```
   1 TTAATTAAGGAGGACCATC ATG AAC GAG GCC ATC GCC GTC GTC GGC ATG TCC TGC CGC CTG CCG   64
   1                     M   N   E   A   I   A   V   V   G   M   S   C   R   L   P    15

  65 AAG GCC TCG AAC CCG GCC GCC TTC TGG GAG CTG CTG CGG AAC GGG GAG AGC GCC GTC ACC  124
  16 K   A   S   N   P   A   A   F   W   E   L   L   R   N   G   E   S   A   V   T    35

 125 GAC GTG CCC TCC GGC CGG TGG ACG TCG GTG CTC GGG GGA GCG GAC GCC GAG GAG CCG GCG  184
  36 D   V   P   S   G   R   W   T   S   V   L   G   G   A   D   A   E   E   P   A    55

 185 GAG TCC GGT GTC CGC CGG GGC GGC TTC CTC GAC TCC CTC GAC CTC TTC GAC GCG GCC TTC  244
  56 E   S   G   V   R   R   G   G   F   L   D   S   L   D   L   F   D   A   A   F    75

 245 TTC GGA ATC TCG CCC CGT GAG GCC GCC GCC ATG GAC CCG CAG CAG CGA CTG GTC CTC GAA  304
  76 F   G   I   S   P   R   E   A   A   M   D   P   Q   Q   R   L   V   L   E    95

 305 CTC GCC TGG GAG GCG CTG GAG GAC GCC GGA ATC GTC CCC GGC ACC CTC GCC GGA AGC CGC  364
  96 L   A   W   E   A   L   E   D   A   G   I   V   P   G   T   L   A   G   S   R   115

 365 ACC GCC GTC TTC GTC GGC ACC CTG CGG GAC GAC TAC ACG AGC CTC CTC TAC CAG CAC GGC  424
 116 T   A   V   F   V   G   T   L   R   D   D   Y   T   S   L   L   Y   Q   H   G   135

 425 GAG CAG GCC ATC ACC CAG CAC ACC ATG GCG GGC GTG AAC CGG GGC GTC ATC GCC AAC CGC  484
 136 E   Q   A   I   T   Q   H   T   M   A   G   V   N   R   G   V   I   A   N   R   155

 485 GTC TCG TAC CAC CTC GGC CTG CAG GGC CCG AGC CTC ACC GTC GAC GCC GCG CAG TCG TCC  544
 156 V   S   Y   H   L   G   L   Q   G   P   S   L   T   V   D   A   A   Q   S   S   175

 545 TCG CTC GTC GCC GTG CAC CTG GCC TGC GAG TCC CTG CGC GCC GGG GAG TCC ACG ACG GCG  604
 176 S   L   V   A   V   H   L   A   C   E   S   L   R   A   G   E   S   T   T   A   195

 605 CTC GTC GCC GGC GTG AAC CTC AAC ATC CTC GCG GAG AGC GCC GTG ACG GAG GAG CGC TTC  664
 196 L   V   A   G   V   N   L   N   I   L   A   E   S   A   V   T   E   E   R   F   215

 665 GGT GGA CTC TCC CCG GAC GGC ACC GCC TAC ACC TTC GAC GCG CGG GCC AAC GGA TTC GTC  724
 216 G   G   L   S   P   D   G   T   A   Y   T   F   D   A   R   A   N   G   F   V   235

 725 CGG GGC GAG GGC GGC GGA GTC GTC GTA CTC AAG CCG CTC TCC CGC GCC CTC GCC GAC GGC  784
 236 R   G   E   G   G   G   V   V   V   L   K   P   L   S   R   A   L   A   D   G   255

 785 GAC CGT GTC CAC GGC GTC ATC CGC GCC AGC GCC GTC AAC AAC GAC GGA GCC ACC CCG GGT  844
 256 D   R   V   H   G   V   I   R   A   S   A   V   N   N   D   G   A   T   P   G   275

 845 CTC ACC GTG CCC AGC AGG GCC GCC CAG GAG AAG GTG CTG CGC GAG GCG TAC CGG AAG GCG  904
 276 L   T   V   P   S   R   A   A   Q   E   K   V   L   R   E   A   Y   R   K   A   295

 905 GCC CTG GAC CCG TCC GCC GTC CAG TAC GTC GAA CTC CAC GGC ACC GGA ACC CCC GTC GGC  964
 296 A   L   D   P   S   A   V   Q   Y   V   E   L   H   G   T   G   T   P   V   G   315

 965 GAC CCC ATC GAG GCC GCC GCG CTC GGC GCC GTC CTC GGC TCG GCG CGC CCC GCG GAC GAA 1024
 316 D   P   I   E   A   A   A   L   G   A   V   L   G   S   A   R   P   A   D   E   335

1025 CCC CTG CTC GTC GGC TCG GCC AAG ACG AAC GTC GGG CAC CTC GAA GGC GCC GCC GGC ATC 1084
 336 P   L   L   V   G   S   A   K   T   N   V   G   H   L   E   G   A   A   G   I   355

1085 GTC GGC CTC ATC AAG ACG CTC CTC GCG CTC GGC CGG CGC CGG ATC CCG GCG AGC CTC AAC 1144
 356 V   G   L   I   K   T   L   L   A   L   G   R   R   R   I   P   A   S   L   N   375

1145 TTC CGT ACG CCC CAC CCG GAC ATC CCG CTC GAC ACC CTC GGG CTC GAC GTG CCC GAC GGC 1204
 376 F   R   T   P   H   P   D   I   P   L   D   T   L   G   L   D   V   P   D   G   395

1205 CTG CGG GAG TGG CCG CAC CCG GAC CGC GAA CTC CTC GCC GGC GTC AGC TCG TTC GGC ATG 1264
 396 L   R   E   W   P   H   P   D   R   E   L   L   A   G   V   S   S   F   G   M   415

1265 GGC GGC ACC AAC GCC CAC GTC GTC CTC AGC GAA GGC CCC GCC CAG GGC GGC GAG CAG CCC 1324
 416 G   G   T   N   A   H   V   V   L   S   E   G   P   A   Q   G   G   E   Q   P   435

1325 GGC ATC GAT GAG GAG ACC CCC GTC GAC AGC GGG GCC GCA CTG CCC TTC GTC GTC ACC GGC 1384
 436 G   I   D   E   E   T   P   V   D   S   G   A   A   L   P   F   V   V   T   G   455

1385 CGC GGC GGC GAG GCC CTG CGC GCC CAG GCC CGG CGC CTG CAC GAG GCC GTC GAA GCG GAC 1444
 456 R   G   G   E   A   L   R   A   Q   A   R   R   L   H   E   A   V   E   A   D   475
```

FIG. 23A

```
1445 CCG GAG CTC GCG CCC GCC GCA CTC GCC CGG TCG CTG GTC ACC ACC CGT ACG GTC TTC ACG 1504
 476 P   E   L   A   P   A   A   L   A   R   S   L   V   T   T   R   T   V   F   T   495

1505 CAC CGG TCG GTC GTC CTC GCC CCG GAC CGC GCC CGC CTC CTC GAC GGC CTC GGC GCC CTC 1564
 496 H   R   S   V   V   L   A   P   D   R   A   R   L   L   D   G   L   G   A   L   515

1565 GCC GCC GGG ACG CCC GCG CCC GGC GTG GTC ACC GGC ACC CCC GCC CCC GGG CGC CTC GCC 1624
 516 A   A   G   T   P   A   P   G   V   V   T   G   T   P   A   P   G   R   L   A   535

1625 GTC CTG TTC AGC GGC CAG GGT GCC CAA CGT ACG GGC ATG GGC ATG GAG TTG TAC GCC GCC 1684
 536 V   L   F   S   G   Q   G   A   Q   R   T   G   M   G   M   E   L   Y   A   A   555

1685 CAC CCC GCC TTC GCG ACG GCC TTC GAC GCC GTC GCC GCC GAA CTG GAC CCC CTC CTC GAC 1744
 556 H   P   A   F   A   T   A   F   D   A   V   A   A   E   L   D   P   L   L   D   575

1745 CGG CCC CTC GCC GAA CTC GTC GCG GCG GGC GAC ACC CTC GAC CGC ACC GTC CAC ACA CAG 1804
 576 R   P   L   A   E   L   V   A   A   G   D   T   L   D   R   T   V   H   T   Q   595

1805 CCC GCG CTC TTC GCC GTG GAG GTC GCC CTC CAC CGC CTC GTC GAG TCC TGG GGC GTC ACG 1864
 596 P   A   L   F   A   V   E   V   A   L   H   R   L   V   E   S   W   G   V   T   615

1865 CCC GAC CTG CTC GCC GGC CAC TCC GTC GGC GAG ATC AGC GCC GCC CAC GTC GCC GGG GTC 1924
 616 P   D   L   L   A   G   H   S   V   G   E   I   S   A   A   H   V   A   G   V   635

1925 CTG TCG CTG CGC GAC GCC GCC CGC CTC GTC GCG GCG CGC GGC CGC CTC ATG CAG GCG CTC 1984
 636 L   S   L   R   D   A   A   R   L   V   A   A   R   G   R   L   M   Q   A   L   655

1985 CCC GAG GGC GGC GCG ATG GTC GCG GTC GAG GCG AGC GAG GAG GAA GTG CTT CCG CAC CTC 2044
 656 P   E   G   G   A   M   V   A   V   E   A   S   E   E   E   V   L   P   H   L   675

2045 GCG GGA CGC GAG CGG GAG CTC TCC CTC GCG GCC GTG AAC GGC CCC CGC GCG GTC GTC CTC 2104
 676 A   G   R   E   R   E   L   S   L   A   A   V   N   G   P   R   A   V   V   L   695

2105 GCG GGC GCC GAG CGC GCC GTC CTC GAC GTC GCC GAG CTG CTG CGC GAA CAG GGC CGC CGG 2164
 696 A   G   A   E   R   A   V   L   D   V   A   E   L   L   R   E   Q   G   R   R   715

2165 ACG AAG CGG CTC AGC GTC TCG CAC GCC TTC CAC TCG CCG CTC ATG GAG CCG ATG CTC GAC 2224
 716 T   K   R   L   S   V   S   H   A   F   H   S   P   L   M   E   P   M   L   D   735

2225 GAC TTC CGC CGG GTC GTC GAA GAG CTG GAC TTC CAG GAG CCC CGC GTC GAC GTC GTG TCC 2284
 736 D   F   R   R   V   V   E   E   L   D   F   Q   E   P   R   V   D   V   V   S   755

2285 ACG GTG ACG GGC CTG CCT GTC ACA GCG GGC CAA TGG ACC GAT CCC GAG TAC TGG GTG GAC 2344
 756 T   V   T   G   L   P   V   T   A   G   Q   W   T   D   P   E   Y   W   V   D   775

2345 CAG GTC CGC AGG CCC GTA CGC TTC CTC GAC GCC GTA CGC ACC CTG GAG GAA TCG GGC GCC 2404
 776 Q   V   R   R   P   V   R   F   L   D   A   V   R   T   L   E   E   S   G   A   795

2405 GAC ACC TTC CTG GAG CTC GGT CCC GAC GGG GTC TGC TCC GCG ATG GCG GCG GAC TCC GTA 2464
 796 D   T   F   L   E   L   G   P   D   G   V   C   S   A   M   A   A   D   S   V   815

2465 CGC GAC CAG GAG GCC GCC ACG GCG GTC TCC GCC CTG CGC AAG GGC CGC CCG GAG CCC CAG 2524
 816 R   D   Q   E   A   A   T   A   V   S   A   L   R   K   G   R   P   E   P   Q   835

2525 TCG CTG CTC GCC GCA CTC ACC ACC GTC TTC GTC CGG GGC CAC GAC GTC GAC TGG ACC GCC 2584
 836 S   L   L   A   A   L   T   T   V   F   V   R   G   H   D   V   D   W   T   A   855

2585 GCG CAC GGG AGC ACC GGC ACG GTC AGG GTG CCC CTG CCG ACC TAC GCC TTC CAG CGC GAA 2644
 856 A   H   G   S   T   G   T   V   R   V   P   L   P   T   Y   A   F   Q   R   E   875

2645 CGC CAC TGG TTC GAC GGC GCC GCG CGA ACG GCG GCC CCG CTC ACG GCG GGC CGA TCG GGC 2704
 876 R   H   W   F   D   G   A   A   R   T   A   A   P   L   T   A   G   R   S   G   895

2705 ACC GGT GCG GGC ACC GGC CCG GCC GCG GGT GTG ACG TCG GGC GAG GGC GAG GGC GAG GGC 2764
 896 T   G   A   G   T   G   P   A   A   G   V   T   S   G   E   G   E   G   E   G   915

2765 GAG GGC GCG GGT GCG GGT GGC GGT GAT CGG CCG GCT CGC CAC GAG ACG ACC GAG CGC GTG 2824
 916 E   G   A   G   A   G   G   G   D   R   P   A   R   H   E   T   T   E   R   V   935

2825 CGC GCA CAC GTC GCC GCC GTC CTC GAG TAC GAC GAC CCG ACC CGC GTC GAA CTC GGC CTC 2884
 936 R   A   H   V   A   A   V   L   E   Y   D   D   P   T   R   V   E   L   G   L   955

2885 ACC TTC AAG GAG CTG GGC TTC GAC TCC CTC ATG TCC GTC GAG CTG CGG AAC GCG CTC GTC 2944
 956 T   F   K   E   L   G   F   D   S   L   M   S   V   E   L   R   N   A   L   V   975

2945 GAC GAC ACG GGA CTG CGC CTG CCC AGC GGA CTG CTC TTC GAC CAC CCG ACG CCG CGC GCC 3004
 976 D   D   T   G   L   R   L   P   S   G   L   L   F   D   H   P   T   P   R   A   995
```

FIG. 23B

```
3005 CTC GCC GCC CAC CTG GGC GAC CTG CTC ACC GGC GGC AGC GGC GAG ACC GGA TCG GCC GAC 3064
 996 L   A   A   H   L   G   D   L   L   T   G   G   S   G   E   T   G   S   A   D  1015

3065 GGG ATA CCG CCC GCG ACC CCG GCG GAC ACC ACC GCC GAG CCC ATC GCG ATC ATC GGC ATG 3124
1016 G   I   P   P   A   T   P   A   D   T   T   A   E   P   I   A   I   I   G   M  1035

3125 GCC TGC CGC TAC CCC GGC GGC GTC ACC TCC CCC GAG GAC CTG TGG CGG CTC GTC GCC GAG 3184
1036 A   C   R   Y   P   G   G   V   T   S   P   E   D   L   W   R   L   V   A   E  1055

3185 GGG CGC GAC GCC GTC TCG GGG CTG CCC ACC GAC CGC GGC TGG GAC GAG GAC CTC TTC GAC 3244
1056 G   R   D   A   V   S   G   L   P   T   D   R   G   W   D   E   D   L   F   D  1075

3245 GCC GAC CCC GAC CGC AGC GGC AAG AGC TCG GTC CGC GAG GGC GGA TTC CTG CAC GAC GCC 3304
1076 A   D   P   D   R   S   G   K   S   S   V   R   E   G   G   F   L   H   D   A  1095

3305 GCC CTG TTC GAC GCC GGC TTC TTC GGG ATA TCG CCC CGC GAG GCC CTC GGC ATG GAC CCG 3364
1096 A   L   F   D   A   G   F   F   G   I   S   P   R   E   A   L   G   M   D   P  1115

3365 CAG CAG CGG CTG CTC CTG GAG ACG GCA TGG GAG GCC GTG GAG CGC GCA GGG CTC GAC CCC 3424
1116 Q   Q   R   L   L   L   E   T   A   W   E   A   V   E   R   A   G   L   D   P  1135

3425 GAA GGC CTC AAG GGC AGC CGG ACG GCC GTC TTC GTC GGC GCC ACC GCC CTG GAC TAC GGC 3484
1136 E   G   L   K   G   S   R   T   A   V   F   V   G   A   T   A   L   D   Y   G  1155

3485 CCG CGC ATG CAC GAC GGC GCC GAG GGC GTC GAG GGC CAC CTC CTG ACC GGG ACC ACG CCC 3544
1156 P   R   M   H   D   G   A   E   G   V   E   G   H   L   L   T   G   T   T   P  1175

3545 AGC GTG ATG TCG GGC CGC ATC GCC TAC CAG CTC GGC CTC ACC GGT CCT GCG GTC ACC GTC 3604
1176 S   V   M   S   G   R   I   A   Y   Q   L   G   L   T   G   P   A   V   T   V  1195

3605 GAC ACG GCC TGC TCG TCC TCG CTC GTC GCG CTG CAC CTG GCC GTC CGT TCG CTG CGG CAG 3664
1196 D   T   A   C   S   S   S   L   V   A   L   H   L   A   V   R   S   L   R   Q  1215

3665 GGC GAG TCG AGC CTC GCG CTC GCC GGC GGA GCG ACC GTC ATG TCG ACA CCG GGC ATG TTC 3724
1216 G   E   S   S   L   A   L   A   G   G   A   T   V   M   S   T   P   G   M   F  1235

3725 GTC GAG TTC TCG CGG CAG CGC GGC CTC GCC GCC GAC GGC CGC TCC AAG GCC TTC TCC GAC 3784
1236 V   E   F   S   R   Q   R   G   L   A   A   D   G   R   S   K   A   F   S   D  1255

3785 TCC GCC GAC GGC ACC TCC TGG GCC GAG GGC GTC GGC CTC CTC GTC GTC GAG CGG CTC TCG 3844
1256 S   A   D   G   T   S   W   A   E   G   V   G   L   L   V   V   E   R   L   S  1275

3845 GAC GCC GAG CGC AAC GGC CAC CCC GTG CTC GCC GTG ATC CGG GGC AGC GCG GTC AAC CAG 3904
1276 D   A   E   R   N   G   H   P   V   L   A   V   I   R   G   S   A   V   N   Q  1295

3905 GAC GGC GCC TCC AAC GGG CTC ACC GCC CCC AAC GGC CCG TCC CAG CAG CGC GTC ATC CGA 3964
1296 D   G   A   S   N   G   L   T   A   P   N   G   P   S   Q   Q   R   V   I   R  1315

3965 CAG GCC CTG GCC GAC GCC GGG CTC ACC CCG GCC GAC GTC GAC GCC GTC GAG GCG CAC GGT 4024
1316 Q   A   L   A   D   A   G   L   T   P   A   D   V   D   A   V   E   A   H   G  1335

4025 ACG GGT ACC CGG CTC GGC GAC CCC ATC GAG GCC GAG GCG ATC CTC GGC ACC TAC GGC CGG 4084
1336 T   G   T   R   L   G   D   P   I   E   A   E   A   I   L   G   T   Y   G   R  1355

4085 GAC CGG GGC GAG GGC GCT CCG CTC CAG CTC GGC TCG CTG AAG TCG AAC ATC GGC CAC GCG 4144
1356 D   R   G   E   G   A   P   L   Q   L   G   S   L   K   S   N   I   G   H   A  1375

4145 CAG GCC GCC GCG GGC GTG GGC GGG CTC ATC AAG ATG GTC CTC GCG ATG CGC CAC GGC GTC 4204
1376 Q   A   A   A   G   V   G   G   L   I   K   M   V   L   A   M   R   H   G   V  1395

4205 CTG CCC AGG ACG CTC CAC GTG GAC CGG CCC ACC ACC CGC GTC GAC TGG GAG GCC GGC GGC 4264
1396 L   P   R   T   L   H   V   D   R   P   T   T   R   V   D   W   E   A   G   G  1415

4265 GTC GAG CTC CTC ACC GAG GAG CGG GAG TGG CCG GAG ACG GGC CGC CCG CGC CGC GCG GCG 4324
1416 V   E   L   L   T   E   E   R   E   W   P   E   T   G   R   P   R   R   A   A  1435

4325 ATC TCC TCC TTC GGC ATC AGC GGC ACC AAC GCC CAC ATC GTG GTC GAA CAG GCC CCG GAA 4384
1436 I   S   S   F   G   I   S   G   T   N   A   H   I   V   V   E   Q   A   P   E  1455

4385 GCC GGG GAG GCG GCG GTC ACC ACC ACC GCC CCG GAA GCA GGG GAA GCC GGG GAA GCG GCG 4444
1456 A   G   E   A   A   V   T   T   T   A   P   E   A   G   E   A   G   E   A   A  1475

4445 GAC ACC ACC GCC ACC ACG ACG CCG GCC GCG GTC GGC GTC CCC GAA CCC GTA CGC GCC CCC 4504
1476 D   T   T   A   T   T   T   P   A   A   V   G   V   P   E   P   V   R   A   P  1495

4505 GTC GTG GTC TCC GCG CGG GAC GCC GCC GCC CTG CGC GCC CAG GCC GTT CGG CTG CGG ACC 4564
1496 V   V   V   S   A   R   D   A   A   A   L   R   A   Q   A   V   R   L   R   T  1515
```

FIG. 23C

```
4565 TTC CTC GAC GGC CGA CCG GAC GTC ACC GTC GCC GAC CTC GGA CGC TCG CTG GCC GCC CGT 4624
1516 F   L   D   G   R   P   D   V   T   V   A   D   L   G   R   S   L   A   A   R   1535

4625 ACC GCC TTC GAG CAC AAG GCC GCC CTC ACC ACC GCC ACC AGG GAC GAG CTG CTC GCC GGG 4684
1536 T   A   F   E   H   K   A   A   L   T   T   A   T   R   D   E   L   L   A   G   1555

4685 CTC GAC GCC CTC GGC CGC GGG GAG CAA GCC ACG GGC CTG GTC ACC GGC GAA CCG GCC AGG 4744
1556 L   D   A   L   G   R   G   E   Q   A   T   G   L   V   T   G   E   P   A   R   1575

4745 GCC GGA CGC ACG GCC TTC CTG TTC ACC GGC CAG GGA GCG CAG CGC GTC GCC ATG GGC GAG 4804
1576 A   G   R   T   A   F   L   F   T   G   Q   G   A   Q   R   V   A   M   G   E   1595

4805 GAA CTG CGC GCC GCG CAC CCC GTG TTC GCC GCC GCC CTC GAC ACC GTG TAC GCG GCC CTC 4864
1596 E   L   R   A   A   H   P   V   F   A   A   A   L   D   T   V   Y   A   A   L   1615

4865 GAC CGT CAC CTC GAC CGG CCG CTG CGG GAG ATC GTC GCC GCC GGG GAG GAG CTG GAC CTC 4924
1616 D   R   H   L   D   R   P   L   R   E   I   V   A   A   G   E   E   L   D   L   1635

4925 ACC GCG TAC ACC CAG CCC GCC CTC TTC GCC TTC GAG GTG GCG CTG TTC CGC CTC CTC GAA 4984
1636 T   A   Y   T   Q   P   A   L   F   A   F   E   V   A   L   F   R   L   L   E   1655

4985 CAC CAC GGC CTC GTC CCC GAC CTG CTC ACC GGC CAC TCC GTC GGC GAG ATC GCC GCC GCG 5044
1656 H   H   G   L   V   P   D   L   L   T   G   H   S   V   G   E   I   A   A   A   1675

5045 CAC GTC GCC GGT GTC CTC TCC CTC GAC GAC GCC GCA CGT CTC GTC ACC GCC CGC GGC CGG 5104
1676 H   V   A   G   V   L   S   L   D   D   A   A   R   L   V   T   A   R   G   R   1695

5105 CTC ATG CAG TCG GCC CGC GAG GGC GGC GCG ATG ATC GCC GTG CAG GCG GGC GAG GCC GAG 5164
1696 L   M   Q   S   A   R   E   G   G   A   M   I   A   V   Q   A   G   E   A   E   1715

5165 GTC GTC GAG TCC CTG AAG GGC TAC GAG GGC AGG GTC GCC GTC GCC GCC GTC AAC GGA CCC 5224
1716 V   V   E   S   L   K   G   Y   E   G   R   V   A   V   A   A   V   N   G   P   1735

5225 ACC GCC GTG GTC GTC TCC GGC GAC GCG GAC GCC GCC GAG GAG ATC CGC GCC GTA TGG GCG 5284
1736 T   A   V   V   V   S   G   D   A   D   A   A   E   E   I   R   A   V   W   A   1755

5285 GGA CGC GGC CGG CGC ACC CGC AGG CTG CGC GTC AGC CAC GCC TTC CAC TCC CCG CAC ATG 5344
1756 G   R   G   R   R   T   R   R   L   R   V   S   H   A   F   H   S   P   H   M   1775

5345 GAC GAC GTC CTC GAC GAG TTC CTC CGG GTC GCC GAG GGC CTG ACC TTC GAG GAG CCG CGG 5404
1776 D   D   V   L   D   E   F   L   R   V   A   E   G   L   T   F   E   E   P   R   1795

5405 ATC CCC GTC GTC TCC ACG GTC ACC GGC GCG CTC GTC ACG TCC GGC GAG CTC ACC TCG CCC 5464
1796 I   P   V   V   S   T   V   T   G   A   L   V   T   S   G   E   L   T   S   P   1815

5465 GCG TAC TGG GTC GAC CAG ATC CGG CGG CCC GTG CGC TTC CTG GAC GCC GTC CGC ACC CTG 5524
1816 A   Y   W   V   D   Q   I   R   R   P   V   R   F   L   D   A   V   R   T   L   1835

5525 GCC GCC CAG GAC GCG ACC GTC CTC GTC GAG ATC GGC CCC GAC GCC GTC CTC ACG GCA CTC 5584
1836 A   A   Q   D   A   T   V   L   V   E   I   G   P   D   A   V   L   T   A   L   1855

5585 GCC GAG GAG GCT CTC GCG CCC GGC ACG GAC GCC CCG GAC GCC CGG GAC GTC ACG GTC GTC 5644
1856 A   E   E   A   L   A   P   G   T   D   A   P   D   A   R   D   V   T   V   V   1875

5645 CCG CTG CTG CGC GCG GGG CGC CCC GAG CCC GAG ACC CTC GCC GCC GGT CTC GCG ACC GCC 5704
1876 P   L   L   R   A   G   R   P   E   P   E   T   L   A   A   G   L   A   T   A   1895

5705 CAT GTC CAC GGC GCA CCC TTG GAC CGG GCG TCG TTC TTC CCG GAC GGG CGC CGC ACG GAC 5764
1896 H   V   H   G   A   P   L   D   R   A   S   F   F   P   D   G   R   R   T   D   1915

5765 CTG CCC ACG TAC GCC TTC CGG CGC GAG CAC TAC TGG CTG ACG CCC GAG GCC CGT ACG GAC 5824
1916 L   P   T   Y   A   F   R   R   E   H   Y   W   L   T   P   E   A   R   T   D   1935

5825 GCC CGC GCA CTC GGC TTC GAC CCG GCG CGG CAC CCG CTG CTG ACG ACC ACG GTC GAG GTC 5884
1936 A   R   A   L   G   F   D   P   A   R   H   P   L   L   T   T   T   V   E   V   1955

5885 GCC GGC GGC GAC GGC GTC CTG CTG ACC GGC CGT CTC TCC CTG ACC GAC CAG CCC TGG CTG 5944
1956 A   G   G   D   G   V   L   L   T   G   R   L   S   L   T   D   Q   P   W   L   1975

5945 GCC GAC CAC ATG GTC AAC GGC GCC GTC CTG TTG CCG GCC ACC GCC TTC CTG GAG CTC GCC 6004
1976 A   D   H   M   V   N   G   A   V   L   L   P   A   T   A   F   L   E   L   A   1995

6005 CTC GCG GCG GGC GAC CAC GTC GGG GCG GTC CGG GTG GAG GAA CTC ACC CTC GAA GCG CCG 6064
1996 L   A   A   G   D   H   V   G   A   V   R   V   E   E   L   T   L   E   A   P   2015

6065 CTC GTC CTG CCC GAG CGG GGC GCC GTC CGC ATC CAG GTC GGC GTG AGC GGC GAC GGC GAG 6124
2016 L   V   L   P   E   R   G   A   V   R   I   Q   V   G   V   S   G   D   G   E   2035
```

FIG. 23D

```
6125 TCG CCG GCC GGG CGC ACC TTC GGT GTG TAC AGC ACC CCC GAC TCC GGC GAC ACC GGT GAC 6184
2036 S   P   A   G   R   T   F   G   V   Y   S   T   P   D   S   G   D   T   G   D   2055

6185 GAC GCG CCC CGG GAG TGG ACC CGC CAT GTC TCC GGC GTA CTC GGC GAA GGG GAC CCG GCC 6244
2056 D   A   P   R   E   W   T   R   H   V   S   G   V   L   G   E   G   D   P   A   2075

6245 ACG GAG TCG GAC CAC CCC GGC ACC GAC GGG GAC GGT TCA GCG GCC TGG CCG CCT GCG GCG 6304
2076 T   E   S   D   H   P   G   T   D   G   D   G   S   A   A   W   P   P   A   A   2095

6305 GCG ACC GCC ACA CCC CTC GAC GGC GTC TAC GAC CGG CTC GCG GAG CTC GGC TAC GGA TAC 6364
2096 A   T   A   T   P   L   D   G   V   Y   D   R   L   A   E   L   G   Y   G   Y   2115

6365 GGT CCG GCC TTC CAG GGC CTG ACG GGG CTG TGG CGC GAC GGC GCC GAC ACG CTC GCC GAG 6424
2116 G   P   A   F   Q   G   L   T   G   L   W   R   D   G   A   D   T   L   A   E   2135

6425 ATC CGG CTG CCC GCG GCG CAG CAC GAG AGC GCG GGG CTC TTC GGC GTA CAC CCG GCG CTG 6484
2136 I   R   L   P   A   A   Q   H   E   S   A   G   L   F   G   V   H   P   A   L   2155

6485 CTC GAC GCG GCG CTC CAC CCG ATC GTC CTG GAG GGC AAC TCA GCT GCC GGT GCC TGT GAC 6544
2156 L   D   A   A   L   H   P   I   V   L   E   G   N   S   A   A   G   A   C   D   2175

6545 GCC GAT ACC GAC GCG ACC GAC CGG ATC CGG CTG CCG TTC GCG TGG GCG GGG GTG ACC CTC 6604
2176 A   D   T   D   A   T   D   R   I   R   L   P   F   A   W   A   G   V   T   L   2195

6605 CAC GCC GAA GGG GCC ACC GCG CTC CGC GTA CGG ATC ACA CCC ACC GGC CCG GAC ACG GTC 6664
2196 H   A   E   G   A   T   A   L   R   V   R   I   T   P   T   G   P   D   T   V   2215

6665 ACG CTC CGC CTC ACC GAC ACC ACC GGT GCG CCC GTG GCC ACC GTG GAG TCC CTG ACC CTG 6724
2216 T   L   R   L   T   D   T   T   G   A   P   V   A   T   V   E   S   L   T   L   2235

6725 CGC GCG GTG GCG AAG GAC CGG CTG GGC ACC ACC GCC GGG CGC GTC GAC GAC GCC CTG TTC 6784
2236 R   A   V   A   K   D   R   L   G   T   T   A   G   R   V   D   D   A   L   F   2255

6785 ACG GTC GTG TGG ACG GAG ACC GGC ACA CCG GAA CCC GCA GGG CGC GGA GCC GTG GAG GTC 6844
2256 T   V   V   W   T   E   T   G   T   P   E   P   A   G   R   G   A   V   E   V   2275

6845 GAG GAA CTC GTC GAC CTC GCC GGC CTC GGC GAC CTC GTG GAG CTC GGC GCC GCG GAC GTC 6904
2276 E   E   L   V   D   L   A   G   L   G   D   L   V   E   L   G   A   A   D   V   2295

6905 GTC CTC CGG GCC GAC CGC TGG ACG CTC GAC GGG GAC CCG TCC GCC GCC GCG CGC ACA GCC 6964
2296 V   L   R   A   D   R   W   T   L   D   G   D   P   S   A   A   A   R   T   A   2315

6965 GTC CGG CGC ACC CTC GCC ATC GTC CAG GAG TTC CTG TCC GAG CCG CGC TTC GAC GGC TCG 7024
2316 V   R   R   T   L   A   I   V   Q   E   F   L   S   E   P   R   F   D   G   S   2335

7025 CGA CTG GTG TGC GTC ACC AGG GGC GCG GTC GCC GCA CTC CCC GGC GAG GAC GTC ACC TCC 7084
2336 R   L   V   C   V   T   R   G   A   V   A   A   L   P   G   E   D   V   T   S   2355

7085 CTC GCC ACC GGC CCC CTC TGG GGC CTC GTC CGC TCC GCC CAG TCC GAG AAC CCG GGA CGC 7144
2356 L   A   T   G   P   L   W   G   L   V   R   S   A   Q   S   E   N   P   G   R   2375

7145 CTG TTC CTC CTG GAC CTG GGT GAA GGC GAA GGC GAG CGC GAC GGA GCC GAG GAG CTG ATC 7204
2376 L   F   L   L   D   L   G   E   G   E   G   E   R   D   G   A   E   E   L   I   2395

7205 CGC GCG GCC ACG GCC GGG GAC GAG CCG CAG CTC GCG GCA CGG GAC GGC CGA CTG CTC GCG 7264
2396 R   A   A   T   A   G   D   E   P   Q   L   A   A   R   D   G   R   L   L   A   2415

7265 CCG AGG CTG GCC CGT ACC GCC GCC CTT TCG AGT GAG GAC ACC GCC GGC GGC GCC GAC CGT 7324
2416 P   R   L   A   R   T   A   A   L   S   S   E   D   T   A   G   G   A   D   R   2435

7325 TTC GGC CCC GAC GGC ACC GTC CTC GTC ACC GGG GGC ACC GGA GGC CTC GGA GCG CTC CTC 7384
2436 F   G   P   D   G   T   V   L   V   T   G   G   T   G   G   L   G   A   L   L   2455

7385 GCC CGC CAC CTC GTG GAG CGT CAC GGG GTG CGC CGG CTG CTG CTG GTG AGC CGC CGC GGG 7444
2456 A   R   H   L   V   E   R   H   G   V   R   R   L   L   L   V   S   R   R   G   2475

7445 GCC GAC GCC CCC GGC GCG GCC GAC CTG GGC GAG GAC CTC GCG GGC CTC GGC GCG GAG GTG 7504
2476 A   D   A   P   G   A   A   D   L   G   E   D   L   A   G   L   G   A   E   V   2495

7505 GCG TTC GCC GCC GCC GAC GCC GCC GAC CGC GAG AGC CTG GCG CGG GCG ATC GCC ACC GTG 7564
2496 A   F   A   A   A   D   A   A   D   R   E   S   L   A   R   A   I   A   T   V   2515

7565 CCC GCC GAG CAT CCG CTG ACG GCC GTC GTG CAC ACG GCG GGA GTC GTC GAC GAC GCG ACG 7624
2516 P   A   E   H   P   L   T   A   V   V   H   T   A   G   V   V   D   D   A   T   2535

7625 GTG GAG GCG CTC ACA CCG GAA CGG CTG GAC GCG GTA CTG CGC CCG AAG GTC GAC GCC GCG 7684
2536 V   E   A   L   T   P   E   R   L   D   A   V   L   R   P   K   V   D   A   A   2555
```

FIG. 23E

```
7685 TGG AAC CTG CAC GAG CTC ACC AAG GAC CTG CGG CTC GAC GCC TTC GTC CTC TTC TCC TCC 7744
2556 W   N   L   H   E   L   T   K   D   L   R   L   D   A   F   V   L   F   S   S   2575

7745 GTC TCC GGC ATC GTC GGC ACC GCC GGC CAG GCC AAC TAC GCG GCG GCC AAC ACG GGC CTC 7804
2576 V   S   G   I   V   G   T   A   G   Q   A   N   Y   A   A   A   N   T   G   L   2595

7805 GAC GCC CTC GCC GCC CAC CGC GCC GCC ACG GGC CTG GCC GCC ACG TCG CTG GCC TGG GGC 7864
2596 D   A   L   A   A   H   R   A   A   T   G   L   A   A   T   S   L   A   W   G   2615

7865 CTC TGG GAC GGC ACG CAC GGC ATG GGC GGC ACG CTC GGC GCC GCC GAC CTC GCC CGC TGG 7924
2616 L   W   D   G   T   H   G   M   G   G   T   L   G   A   A   D   L   A   R   W   2635

7925 AGC CGG GCC GGA ATC ACC CCG CTC ACC CCG CTG CAG GGC CTC GCG CTC TTC GAC GCC GCG 7984
2636 S   R   A   G   I   T   P   L   T   P   L   Q   G   L   A   L   F   D   A   A   2655

7985 GTC GCC AGG GAC GAC GCC CTC CTC GTA CCC GCC GGG CTC CGT CCC ACC GCC CAC CGG GGC 8044
2656 V   A   R   D   D   A   L   L   V   P   A   G   L   R   P   T   A   H   R   G   2675

8045 ACG GAC GGA CAG CCT CCT GCG CTG TGG CGC GGC CTC GTC CGG GCG CGC CCG CGC CGT GCC 8104
2676 T   D   G   Q   P   P   A   L   W   R   G   L   V   R   A   R   P   R   R   A   2695

8105 GCG CGG ACG GCC GCC GAG GCG GCG GAC ACG ACC GGC GGC TGG CTG AGC GGG CTC GCC GCA 8164
2696 A   R   T   A   A   E   A   A   D   T   T   G   G   W   L   S   G   L   A   A   2715

8165 CAG TCC CCC GAG GAG CGG CGC AGC ACA GCC GTC ACG CTC GTG ACG GGT GTC GTC GCG GAC 8224
2716 Q   S   P   E   E   R   R   S   T   A   V   T   L   V   T   G   V   V   A   D   2735

8225 GTC CTC GGG CAC GCC GAC TCC GCC GCG GTC GGG GCG GAG CGG TCC TTC AAG GAC CTC GGC 8284
2736 V   L   G   H   A   D   S   A   A   V   G   A   E   R   S   F   K   D   L   G   2755

8285 TTC GAC TCC CTG GCC GGG GTG GAG CTC CGC AAC CGG CTG AAC GCC GCC ACC GGC CTG CGG 8344
2756 F   D   S   L   A   G   V   E   L   R   N   R   L   N   A   A   T   G   L   R   2775

8345 CTC CCC GCG ACC ACG GTC TTC GAC CAT CCC TCG CCG GCC GCG CTC GCG TCC CAT CTC CTC 8404
2776 L   P   A   T   T   V   F   D   H   P   S   P   A   A   L   A   S   H   L   L   2795

8405 GCC CAG GTG CCC GGG TTG AAG GAG GGG ACG GCG GCG ACC GCG ACC GTC GTG GCC GAG CGG 8464
2796 A   Q   V   P   G   L   K   E   G   T   A   A   T   A   T   V   V   A   E   R   2815

8465 GGC GCT TCC TTC GGT GAC CGT GCG ACC GAC GAC GAT CCG ATC GCG ATC GTG GGC ATG GCA 8524
2816 G   A   S   F   G   D   R   A   T   D   D   D   P   I   A   I   V   G   M   A   2835

8525 TGC CGC TAT CCG GGT GGT GTG TCG TCG CCG GAG GAC CTG TGG CGG CTG GTG GCC GAG GGG 8584
2836 C   R   Y   P   G   G   V   S   S   P   E   D   L   W   R   L   V   A   E   G   2855

8585 ACG GAC GCG ATC AGC GAG TTC CCC GTC AAC CGC GGC TGG GAC CTG GAG AGC CTC TAC GAC 8644
2856 T   D   A   I   S   E   F   P   V   N   R   G   W   D   L   E   S   L   Y   D   2875

8645 CCG GAT CCC GAG TCG AAG GGC ACC ACG TAC TGC CGG GAG GGC GGG TTC CTG GAA GGC GCC 8704
2876 P   D   P   E   S   K   G   T   T   Y   C   R   E   G   G   F   L   E   G   A   2895

8705 GGT GAC TTC GAC GCC GCC TTC TTC GGC ATC TCG CCG CGC GAG GCC CTG GTG ATG GAC CCG 8764
2896 G   D   F   D   A   A   F   F   G   I   S   P   R   E   A   L   V   M   D   P   2915

8765 CAG CAG CGG CTG CTG CTG GAG GTG TCC TGG GAG GCG CTG GAA CGC GCG GGC ATC GAC CCG 8824
2916 Q   Q   R   L   L   L   E   V   S   W   E   A   L   E   R   A   G   I   D   P   2935

8825 TCC TCG CTG CGC GGC AGC CGC GGT GGT GTC TAC GTG GGC GCC GCG CAC GGC TCG TAC GCC 8884
2936 S   S   L   R   G   S   R   G   G   V   Y   V   G   A   A   H   G   S   Y   A   2955

8885 TCC GAT CCC CGG CTG GTG CCC GAG GGC TCG GAG GGC TAT CTG CTG ACC GGC AGC GCC GAC 8944
2956 S   D   P   R   L   V   P   E   G   S   E   G   Y   L   L   T   G   S   A   D   2975

8945 GCG GTG ATG TCC GGC CGC ATC TCC TAC GCG CTC GGT CTC GAA GGA CCG TCC ATG ACG GTG 9004
2976 A   V   M   S   G   R   I   S   Y   A   L   G   L   E   G   P   S   M   T   V   2995

9005 GAG ACG GCC TGC TCC TCC TCG CTG GTG GCG CTG CAT CTG GCG GTA CGG GCG CTG CGG CAC 9064
2996 E   T   A   C   S   S   S   L   V   A   L   H   L   A   V   R   A   L   R   H   3015

9065 GGC GAG TGC GGG CTC GCG CTG GCG GGC GGG GTG GCG GTG ATG GCC GAT CCG GCG GCG TTC 9124
3016 G   E   C   G   L   A   L   A   G   G   V   A   V   M   A   D   P   A   A   F   3035

9125 GTG GAG TTC TCC CGG CAG AAG GGG CTG GCC GCC GAC GGC CGC TGC AAG GCG TTC TCG GCC 9184
3036 V   E   F   S   R   Q   K   G   L   A   A   D   G   R   C   K   A   F   S   A   3055

9185 GCC GCC GAC GGC ACC GGC TGG GCC GAG GGC GTC GGC GTG CTC GTC CTG GAG CGG CTG TCG 9244
3056 A   A   D   G   T   G   W   A   E   G   V   G   V   L   V   L   E   R   L   S   3075
```

FIG. 23F

```
9245 GAC GCG CGC CGC GCG GGG CAC ACG GTC CTC GGC CTG GTC ACC GGC ACC GCG GTC AAC CAG  9304
3076 D   A   R   R   A   G   H   T   V   L   G   L   V   T   G   T   A   V   N   Q    3095

9305 GAC GGT GCC TCC AAC GGG CTG ACC GCG CCC AAC GGC CCA GCC CAG CAA CGC GTC ATC GCC  9364
3096 D   G   A   S   N   G   L   T   A   P   N   G   P   A   Q   Q   R   V   I   A    3115

9365 GAG GCG CTC GCC GAC GCC GGG CTG TCC CCG GAG GAC GTG GAC GCG GTC GAG GCG CAC GGC  9424
3116 E   A   L   A   D   A   G   L   S   P   E   D   V   D   A   V   E   A   H   G    3135

9425 ACC GGC ACC CGG CTC GGC GAC CCC ATC GAG GCC GGG GCG CTG CTC GCC GCC TCC GGA CGG  9484
3136 T   G   T   R   L   G   D   P   I   E   A   G   A   L   L   A   A   S   G   R    3155

9485 AAC CGT TCC GGC GAC CAC CCG CTG TGG CTC GGC TCG CTG AAG TCC AAC ATC GGG CAT GCC  9544
3156 N   R   S   G   D   H   P   L   W   L   G   S   L   K   S   N   I   G   H   A    3175

9545 CAG GCC GCC GCC GGT GTC GGC GGC GTC ATC AAG ATG CTC CAG GCG CTG CGG CAC GGC TTG  9604
3176 Q   A   A   A   G   V   G   G   V   I   K   M   L   Q   A   L   R   H   G   L    3195

9605 CTG CCC CGC ACC CTC CAC GCC GAC GAG CCG ACC CCG CAT GCC GAC TGG AGC TCC GGC CGG  9664
3196 L   P   R   T   L   H   A   D   E   P   T   P   H   A   D   W   S   S   G   R    3215

9665 GTA CGG CTG CTC ACC TCC GAG GTG CCG TGG CAG CGG ACC GGC CGG CCC CGG CGG ACC GGG  9724
3216 V   R   L   L   T   S   E   V   P   W   Q   R   T   G   R   P   R   R   T   G    3235

9725 GTG TCC GCC TTC GGC GTC GGC GGC ACC AAT GCC CAT GTC GTC CTC GAA GAG GCA CCC GCC  9784
3236 V   S   A   F   G   V   G   G   T   N   A   H   V   V   L   E   E   A   P   A    3255

9785 CCG CCC GCG CCG GAA CCG GCC GGG GAG GCC CCC GGC GGC TCC CGC GCC GCA GAA GGG GCG  9844
3256 P   P   A   P   E   P   A   G   E   A   P   G   G   S   R   A   A   E   G   A    3275

9845 GAA GGG CCC CTG GCC TGG GTG GTC TCC GGA CGC GAC GAG CCG GCC CTG CGG TCC CAG GCC  9904
3276 E   G   P   L   A   W   V   V   S   G   R   D   E   P   A   L   R   S   Q   A    3295

9905 CGG CGG CTC CGC GAC CAC CTC TCC CGC ACC CCC GGG GCC CGC CCG CGT GAC ATC GCC TTC  9964
3296 R   R   L   R   D   H   L   S   R   T   P   G   A   R   P   R   D   I   A   F    3315

9965 TCC CTC GCC GCC ACG CGC GCA GCC TTT GAC CAC CGC GCC GTG CTG ATC GGC TCG GAC GGG 10024
3316 S   L   A   A   T   R   A   A   F   D   H   R   A   V   L   I   G   S   D   G    3335

10025 GCC GAA CTC GCC GCC GCC CTG GAC GCG TTG GCC GAA GGA CGC GAC GGT CCG GCG GTG GTG 10084
 3336 A   E   L   A   A   A   L   D   A   L   A   E   G   R   D   G   P   A   V   V    3355

10085 CGC GGA GTC CGC GAC CGG GAC GGC AGG ATG GCC TTC CTC TTC ACC GGG CAG GGC AGC CAG 10144
 3356 R   G   V   R   D   R   D   G   R   M   A   F   L   F   T   G   Q   G   S   Q    3375

10145 CGC GCC GGG ATG GCC CAC GAC CTG CAT GCC GCC CAT ACC TTC TTC GCG TCC GCC CTC GAC 10204
 3376 R   A   G   M   A   H   D   L   H   A   A   H   T   F   F   A   S   A   L   D    3395

10205 GAG GTG ACG GAC CGT CTC GAC CCG CTG CTC GGC CGG CCG CTC GGC GCG CTG CTG GAC GCC 10264
 3396 E   V   T   D   R   L   D   P   L   L   G   R   P   L   G   A   L   L   D   A    3415

10265 CGA CCC GGC TCG CCC GAA GCG GCA CTC CTG GAC CGG ACC GAG TAC ACC CAG CCG GCG CTC 10324
 3416 R   P   G   S   P   E   A   A   L   L   D   R   T   E   Y   T   Q   P   A   L    3435

10325 TTC GCC GTC GAG GTG GCG CTC CAC CGG CTG CTG GAG CAC TGG GGG ATG CGC CCC GAC CTG 10384
 3436 F   A   V   E   V   A   L   H   R   L   L   E   H   W   G   M   R   P   D   L    3455

10385 CTG CTG GGG CAC TCG GTG GGC GAA CTG GCG GCC GCC CAC GTC GCG GGT GTG CTC GAT CTC 10444
 3456 L   L   G   H   S   V   G   E   L   A   A   A   H   V   A   G   V   L   D   L    3475

10445 CAC GAC GCC TGC GCG CTG GTG GCC GCC CGC GGC AGG CTG ATG CAG CGC CTG CCG CCC GGC 10504
 3476 D   D   A   C   A   L   V   A   A   R   G   R   L   M   Q   R   L   P   P   G    3495

10505 GGC GCG ATG GTC TCC GTG CGG GCC GGC GAG GAC GAG GTC CGC GCA CTG CTG GCC GGC CGC 10564
 3496 G   A   M   V   S   V   R   A   G   E   D   E   V   R   A   L   L   A   G   R    3515

10565 GAG GAC GCC GTC TGC GTC GCC GCG GTG AAC GGC CCC CGG TCG GTG GTG ATC TCC GGC GCG 10624
 3516 E   D   A   V   C   V   A   A   V   N   G   P   R   S   V   V   I   S   G   A    3535

10625 GAG GAA GCG GTG GCC GAG GCG GCG GCG CAG CTC GCC GGA CGA GGC CGC CGC ACC AGG CGG 10684
 3536 E   E   A   V   A   E   A   A   A   Q   L   A   G   R   G   R   R   T   R   R    3555

10685 CTC CGC GTC GCG CAC GCC TTC CAC TCA CCC CTG ATG GAC GGC ATG CTC GCC GGA TTC CGG 10744
 3556 L   R   V   A   H   A   F   H   S   P   L   M   D   G   M   L   A   G   F   R    3575

10745 GAG GTC GCC GCC GGC CTG CGC TAC CGG GAA CCG GAG CTG ACG GTC GTC TCC ACG GTC ACG 10804
 3576 E   V   A   A   G   L   R   Y   R   E   P   E   L   T   V   V   S   T   V   T    3595
```

FIG. 23G

```
10805 GGG CGG CCC GCC CGC CCC GGT GAA CTC ACC GGC CCC GAC TAC TGG GTG GCC CAG GTC CGT 10864
 3596 G   R   P   A   R   P   G   E   L   T   G   P   D   Y   W   V   A   Q   V   R   3615

10865 GAG CCC GTG CGC TTC GCG GAC GCG GTC CGC ACG GCA CAC CGC CTC GGA GCC CGC ACC TTC 10924
 3616 E   P   V   R   F   A   D   A   V   R   T   A   H   R   L   G   A   R   T   F   3635

10925 CTG GAG ACC GGC CCG GAC GGC GTG CTG TGC GGC ATG GCA GAG GAG TGC CTG GAG GAC GAC 10984
 3636 L   E   T   G   P   D   G   V   L   C   G   M   A   E   E   C   L   E   D   D   3655

10985 ACC GTG GCC CTG CTG CCG GCG ATC CAC AAG CCC GGC ACC GCG CCG CAC GGT CCG GCG GCT 11044
 3656 T   V   A   L   L   P   A   I   H   K   P   G   T   A   P   H   G   P   A   A   3675

11045 CCC GGC GCG CTG CGG GCG GCC GCC GCC GCG TAC GGC CGG GGC GCC CGG GTG GAC TGG GCC 11104
 3676 P   G   A   L   R   A   A   A   A   Y   G   R   G   A   R   V   D   W   A   3695

11105 GGG ATG CAC GCC GAC GGC CCC GAG GGG CCG GCC CGC CGC GTC GAA CTG CCC GTC CAC GCC 11164
 3696 G   M   H   A   D   G   P   E   G   P   A   R   R   V   E   L   P   V   H   A   3715

11165 TTC CGG CAC CGC CGC TAC TGG CTC GCC CCG GGC CGC GCG GCG GAC ACC GAC GAC TGG ATG 11224
 3716 F   R   H   R   R   Y   W   L   A   P   G   R   A   A   D   T   D   D   W   M   3735

11225 TAC CGG ATC GGC TGG GAC CGG CTG CCG GCT GTG ACC GGC GGG GCC CGG ACC GCC GGC CGC 11284
 3736 Y   R   I   G   W   D   R   L   P   A   V   T   G   G   A   R   T   A   G   R   3755

11285 TGG CTG GTG ATC CAC CCC GAC AGC CCG CGC TGC CGG GAG CTG TCC GGC CAC GCC GAA CGC 11344
 3756 W   L   V   I   H   P   D   S   P   R   C   R   E   L   S   G   H   A   E   R   3775

11345 GCG CTG CGC GCC GCG GGC GCG AGC CCC GTA CCG CTG CCC GTG GAC GCT CCG GCC GCC GAC 11404
 3776 A   L   R   A   A   G   A   S   P   V   P   L   P   V   D   A   P   A   A   D   3795

11405 CGG GCG TCC TTC GCG GCA CTG CTG CGC TCC GCC ACC GGA CCT GAC ACA CGA GGT GAC ACA 11464
 3796 R   A   S   F   A   A   L   L   R   S   A   T   G   P   D   T   R   G   D   T   3815

11465 GCC GCG CCC GTG GCC GGT GTG CTG TCG CTG CTG TCC GAG GAG GAT CGG CCC CAT CGC CAG 11524
 3816 A   A   P   V   A   G   V   L   S   L   L   S   E   E   D   R   P   H   R   Q   3835

11525 CAC GCC CCG GTA CCC GCC GGG GTC CTG GCG ACG CTG TCC CTG ATG CAG GCT ATG GAG GAG 11584
 3836 H   A   P   V   P   A   G   V   L   A   T   L   S   L   M   Q   A   M   E   E   3855

11585 GAG GCG GTG GAG GCT CGC GTG TGG TGC GTC TCC CGC GCC GCG GTC GCC GCC GCC GAC CGG 11644
 3856 E   A   V   E   A   R   V   W   C   V   S   R   A   A   V   A   A   A   D   R   3875

11645 GAA CGG CCC GTC GGC GCG GGC GCC GCC CTG TGG GGG CTG GGG CGG GTG GCC GCC CTG GAA 11704
 3876 E   R   P   V   G   A   G   A   A   L   W   G   L   G   R   V   A   A   L   E   3895

11705 CGC CCC ACC CGG TGG GGC GGT CTC GTG GAC CTG CCC GCC TCG CCC GGT GCG GCG CAC TGG 11764
 3896 R   P   T   R   W   G   G   L   V   D   L   P   A   S   P   G   A   A   H   W   3915

11765 GCG GCC GCC GTG GAA CGG CTC GCC GGT CCC GAG GAC CAG ATC GCC GTG CGC GCG TCC GGC 11824
 3916 A   A   A   V   E   R   L   A   G   P   E   D   Q   I   A   V   R   A   S   G   3935

11825 AGT TGG GGC CGG CGC CTC ACC AGG CTG CCG CGC GAC GGC GGC GGC CGG ACG GCC GCA CCC 11884
 3936 S   W   G   R   R   L   T   R   L   P   R   D   G   G   G   R   T   A   A   P   3955

11885 GCG TAC CGG CCG CGC GGC ACG GTG CTC GTC ACC GGT GGC ACC GGC GCG CTC GGC GGG CAT 11944
 3956 A   Y   R   P   R   G   T   V   L   V   T   G   G   T   G   A   L   G   G   H   3975

11945 CTC GCC CGC TGG CTC GCC GCG GCG GGC GCC GAA CAC CTG GCG CTC ACC AGC CGC CGG GGC 12004
 3976 L   A   R   W   L   A   A   A   G   A   E   H   L   A   L   T   S   R   R   G   3995

12005 CCG GAC GCG CCC GGC GCC GCC GGA CTC GAG GCC GAA CTC CTC CTC CTG GGC GCC AAG GTG 12064
 3996 P   D   A   P   G   A   A   G   L   E   A   E   L   L   L   L   G   A   K   V   4015

12065 ACG TTC GCC GCC TGC GAC ACC GCC GAC CGC GAC GGC CTC GCC CGG GTC CTG CGG GCG ATA 12124
 4016 T   F   A   A   C   D   T   A   D   R   D   G   L   A   R   V   L   R   A   I   4035

12125 CCG GAG GAC ACC CCG CTC ACC GCG GTG TTC CAC GCC GCG GGC GTA CCG CAG GTC ACG CCG 12184
 4036 P   E   D   T   P   L   T   A   V   F   H   A   A   G   V   P   Q   V   T   P   4055

12185 CTG TCC CGT ACC TCG CCC GAG CAC TTC GCC GAC GTG TAC GCG GGC AAG GCG GCG GGC GCC 12244
 4056 L   S   R   T   S   P   E   H   F   A   D   V   Y   A   G   K   A   A   G   A   4075

12245 GCG CAC CTG GAC GAA CTG ACC CGC GAA CTC GGC GCC GGA CTC GAC GCG TTC GTC CTC TAC 12304
 4076 A   H   L   D   E   L   T   R   E   L   G   A   G   L   D   A   F   V   L   Y   4095

12305 TCC TCC GGC GCC GGC GTC TGG GGC AGC GCC GGC CAG GGT GCC TAC GCC GCC GCC AAC GCC 12364
 4096 S   S   G   A   G   V   W   G   S   A   G   Q   G   A   Y   A   A   A   N   A   4115
```

FIG. 23H

```
12365 GCC CTG GAC GCG CTC GCC CGG CGC CGT GCG GCG GAC GGA CTC CCC GCC ACC TCC ATC GCC 12424
 4116 A   L   D   A   L   A   R   R   R   A   A   D   G   L   P   A   T   S   I   A   4135

12425 TGG GGC GTG TGG GGC GGC GGC GGT ATG GGG GCC GAC GAG GCG GGC GCG GAG TAT CTG GGC 12484
 4136 W   G   V   W   G   G   G   G   M   G   A   D   E   A   G   A   E   Y   L   G   4155

12485 CGG CGC GGT ATG CGC CCC ATG GCA CCG GTC TCC GCG CTC CGG GCG ATG GCC ACC GCC ATC 12544
 4156 R   R   G   M   R   P   M   A   P   V   S   A   L   R   A   M   A   T   A   I   4175

12545 GCC TCC GGG GAA CCC TGC CCC ACC GTC ACC CAC ACC GAC TGG GAG CGC TTC GGC GAG GGC 12604
 4176 A   S   G   E   P   C   P   T   V   T   H   T   D   W   E   R   F   G   E   G   4195

12605 TTC ACC GCC TTC CGG CCC AGC CCT CTG ATC GCG GGG CTC GGC ACG CCG GGC GGC GGC CGG 12664
 4196 F   T   A   F   R   P   S   P   L   I   A   G   L   G   T   P   G   G   G   R   4215

12665 GCG GCG GAG ACC CCC GAG GAG GGG AAC GCC ACC GCT GCG GCG GAC CTC ACC GCC CTG CCG 12724
 4216 A   A   E   T   P   E   E   G   N   A   T   A   A   A   D   L   T   A   L   P   4235

12725 CCC GCC GAA CTC CGC ACC GCG CTG CGC GAG CTG GTG CGA GCC CGG ACC GCC GCG GCG CTC 12784
 4236 P   A   E   L   R   T   A   L   R   E   L   V   R   A   R   T   A   A   A   L   4255

12785 GGC CTC GAC GAC CCG GCC GAG GTC GCC GAG GGC GAA CGG TTC CCC GCC ATG GGC TTC GAC 12844
 4256 G   L   D   D   P   A   E   V   A   E   G   E   R   F   P   A   M   G   F   D   4275

12845 TCC CTG GCC ACC GTA CGG CTG CGC CGC GGA CTC GCC TCG GCC ACG GGC CTC GAC CTG CCC 12904
 4276 S   L   A   T   V   R   L   R   R   G   L   A   S   A   T   G   L   D   L   P   4295

12905 CCC GAT CTG CTC TTC GAC CGG GAC ACC CCG GCC GCG CTC GCC GCC CAC CTG GCC GAA CTG 12964
 4296 P   D   L   L   F   D   R   D   T   P   A   A   L   A   A   H   L   A   E   L   4315

12965 CTC GCC ACC GCA CGG GAC CAC GGA CCC GGC GGC CCC GGG ACC GGT GCC GCG CCG GCC GAT 13024
 4316 L   A   T   A   R   D   H   G   P   G   G   P   G   T   G   A   A   P   A   D   4335

13025 GCC GGA AGC GGC CTG CCG GCC CTC TAC CGG GAG GCC GTC CGC ACC GGC CGG GCC GCG GAA 13084
 4336 A   G   S   G   L   P   A   L   Y   R   E   A   V   R   T   G   R   A   A   E   4355

13085 ATG GCC GAA CTG CTC GCC GCC GCT TCC CGG TTC CGC CCC GCC TTC GGG ACG GCG GAC CGG 13144
 4356 M   A   E   L   L   A   A   A   S   R   F   R   P   A   F   G   T   A   D   R   4375

13145 CAG CCG GTG GCC CTC GTG CCG CTG GCC GAC GGC GCG GAG GAC ACC GGG CTC CCG CTG CTC 13204
 4376 Q   P   V   A   L   V   P   L   A   D   G   A   E   D   T   G   L   P   L   L   4395

13205 GTG GGC TGC GCC GGG ACG GCG GTG GCC TCC GGC CCG GTG GAG TTC ACC GCC TTC GCC GGA 13264
 4396 V   G   C   A   G   T   A   V   A   S   G   P   V   E   F   T   A   F   A   G   4415

13265 GCG CTG GCG GAC CTC CCG GCG GCG GCC CCG ATG GCC GCG CTG CCG CAG CCC GGC TTT CTG 13324
 4416 A   L   A   D   L   P   A   A   A   P   M   A   A   L   P   Q   P   G   F   L   4435

13325 CCG GGA GAA CGA GTC CCG GCC ACC CCG GAG GCA TTG TTC GAG GCC CAG GCG GAA GCG CTG 13384
 4436 P   G   E   R   V   P   A   T   P   E   A   L   F   E   A   Q   A   E   A   L   4455

13385 CTG CGC TAC GCG GCC GGC CGG CCC TTC GTG CTG CTG GGG CAC TCC GCC GGC GCC AAC ATG 13444
 4456 L   R   Y   A   A   G   R   P   F   V   L   L   G   H   S   A   G   A   N   M   4475

13445 GCC CAC GCC CTG ACC CGT CAT CTG GAG GCG AAC GGT GGC GGC CCC GCA GGG CTG GTG CTC 13504
 4476 A   H   A   L   T   R   H   L   E   A   N   G   G   G   P   A   G   L   V   L   4495

13505 ATG GAC ATC TAC ACC CCC GCC GAC CCC GGC GCG ATG GGC GTC TGG CGG AAC GAC ATG TTC 13564
 4496 M   D   I   Y   T   P   A   D   P   G   A   M   G   V   W   R   N   D   M   F   4515

13565 CAG TGG GTC TGG CGG CGC TCG GAC ATC CCC CCG GAC GAC CAC CGC CTC ACG GCC ATG GGC 13624
 4516 Q   W   V   W   R   R   S   D   I   P   P   D   D   H   R   L   T   A   M   G   4535

13625 GCC TAC CAC CGG CTG CTT CTC GAC TGG TCG CCC ACC CCC GTC CGC GCC CCC GTA CTG CAT 13684
 4536 A   Y   H   R   L   L   L   D   W   S   P   T   P   V   R   A   P   V   L   H   4555

13685 CTG CGC GCC GCG GAA CCC ATG GGC GAC TGG CCA CCC GGG GAC ACC GGC TGG CAG TCC CAC 13744
 4556 L   R   A   A   E   P   M   G   D   W   P   P   G   D   T   G   W   Q   S   H   4575

13745 TGG GAC GGC GCG CAC ACC ACC GCC GGC ATC CCC GGA AAC CAC TTC ACG ATG ATG ACC GAA 13804
 4576 W   D   G   A   H   T   T   A   G   I   P   G   N   H   F   T   M   M   T   E   4595

13805 CAC GCC TCC GCC GCC GCC CGG CTC GTG CAC GGC TGG CTC GCG GAA CGG ACC CCG TCC GGG 13864
 4596 H   A   S   A   A   A   R   L   V   H   G   W   L   A   E   R   T   P   S   G   4615

13865 CAG GGC GGG TCA CCG TCC CGC GCG GCG GGG AGA GAG GAG AGG CCG TGA ACACGGCAGCCGGCCC 13928
 4616 Q   G   G   S   P   S   R   A   A   G   R   E   E   R   P   *                   4631
```

FIG. 231

```
13929 GACCGGCACCGCCGCCGGCGGCACCACCGCCCCGGCGGCGGCACACGACCTGTCCCGCGCCGGACGCAGGCTCCAACTCA 14008

14009 CCCGGGCCGCACAGTGGTTCGCCGGCAACCAGGGAGACCCCTACGGG ATG ATC CTG CGC GCC GGC ACC GCC  14079
    1                                                M   I   L   R   A   G   T   A    8

14080 GAC CCG GCA CCG TAC GAG GAA GAG ATC CCC GGG TAC CGA GCT CGA ATT CTT AAT TAA GGAG 14140
    9 D   P   A   P   Y   E   E   E   I   P   G   Y   R   A   R   I   L   N   *       27

14141 GTCGTAG ATG AGT AAC AAG AAC AAC GAT GAG CTG CAG CGG CAG GCC TCG GAA AAC ACC CTG  14201
    1        M   S   N   K   N   N   D   E   L   Q   R   Q   A   S   E   N   T   L   18

14202 GGG CTG AAC CCG GTC ATC GGT ATC CGC CGC AAA GAC CTG TTG AGC TCG GCA CGC ACC GTG  14261
   19 G   L   N   P   V   I   G   I   R   R   K   D   L   L   S   S   A   R   T   V   38

14262 CTG CGC CAG GCC GTG CGC CAA CCG CTG CAC AGC GCC AAG CAT GTG GCC CAC TTT GGC CTG  14321
   39 L   R   Q   A   V   R   Q   P   L   H   S   A   K   H   V   A   H   F   G   L   58

14322 GAG CTG AAG AAC GTG CTG CTG GGC AAG TCC AGC CTT GCC CCG GAA AGC GAC GAC CGT CGC  14381
   59 E   L   K   N   V   L   L   G   K   S   S   L   A   P   E   S   D   D   R   R   78

14382 TTC AAT GAC CCG GCA TGG AGC AAC AAC CCA CTT TAC CGC CGC TAC CTG CAA ACC TAT CTG  14441
   79 F   N   D   P   A   W   S   N   N   P   L   Y   R   R   Y   L   Q   T   Y   L   98

14442 GCC TGG CGC AAG GAG CTG CAG GAC TGG ATC GGC AAC AGC GAC CTG TCG CCC CAG GAC ATC  14501
   99 A   W   R   K   E   L   Q   D   W   I   G   N   S   D   L   S   P   Q   D   I  118

14502 AGC CGC GGC CAG TTC GTC ATC AAC CTG ATG ACC GAA GCC ATG GCT CCG ACC AAC ACC CTG  14561
  119 S   R   G   Q   F   V   I   N   L   M   T   E   A   M   A   P   T   N   T   L  138

14562 TCC AAC CCG GCA GCA GTC AAA CGC TTC TTC GAA ACC GGC GGC AAG AGC CTG CTC GAT GGC  14621
  139 S   N   P   A   A   V   K   R   F   F   E   T   G   G   K   S   L   L   D   G  158

14622 CTG TCC AAC CTG GCC AAG GAC CTG GTC AAC AAC GGT GGC ATG CCC AGC CAG GTG AAC ATG  14681
  159 L   S   N   L   A   K   D   L   V   N   N   G   G   M   P   S   Q   V   N   M  178

14682 GAC GCC TTC GAG GTG GGC AAG AAC CTG GGC ACC AGT GAA GGC GCC GTG GTG TAC CGC AAC  14741
  179 D   A   F   E   V   G   K   N   L   G   T   S   E   G   A   V   V   Y   R   N  198

14742 GAT GTG CTG GAG CTG ATC CAG TAC AAG CCC ATC ACC GAG CAG GTG CAT GCC CGC CCG CTG  14801
  199 D   V   L   E   L   I   Q   Y   K   P   I   T   E   Q   V   H   A   R   P   L  218

14802 CTG GTG GTG CCG CCG CAG ATC AAC AAG TTC TAC GTA TTC GAC CTG AGC CCG GAA AAG AGC  14861
  219 L   V   V   P   P   Q   I   N   K   F   Y   V   F   D   L   S   P   E   K   S  238

14862 CTG GCA CGC TAC TGC CTG CGC TCG CAG CAG CAG ACC TTC ATC ATC AGC TGG CGC AAC CCG  14921
  239 L   A   R   Y   C   L   R   S   Q   Q   Q   T   F   I   I   S   W   R   N   P  258

14922 ACC AAA GCC CAG CGC GAA TGG GGC CTG TCC ACC TAC ATC GAC GCG CTC AAG GAG GCG GTC  14981
  259 T   K   A   Q   R   E   W   G   L   S   T   Y   I   D   A   L   K   E   A   V  278

14982 GAC GCG GTG CTG GCG ATT ACC GGC AGC AAG GAC CTG AAC ATG CTC GGT GCC TGC TCC GGC  15041
  279 D   A   V   L   A   I   T   G   S   K   D   L   N   M   L   G   A   C   S   G  298

15042 GGC ATC ACC TGC ACG GCA TTG GTC GGC CAC TAT GCC GCC CTC GGC GAA AAC AAG GTC AAT  15101
  299 G   I   T   C   T   A   L   V   G   H   Y   A   A   L   G   E   N   K   V   N  318

15102 GCC CTG ACC CTG CTG GTC AGC GTG CTG GAC ACC ACC ATG GAC AAC CAG GTC GCC CTG TTC  15161
  319 A   L   T   L   L   V   S   V   L   D   T   T   M   D   N   Q   V   A   L   F  338

15162 GTC GAC GAG CAG ACT TTG GAG GCC GCC AAG CGC CAC TCC TAC CAG GCC GGT GTG CTC GAA  15221
  339 V   D   E   Q   T   L   E   A   A   K   R   H   S   Y   Q   A   G   V   L   E  358

15222 GGC AGC GAG ATG GCC AAG GTG TTC GCC TGG ATG CGC CCC AAC GAC CTG ATC TGG AAC TAC  15281
  359 G   S   E   M   A   K   V   F   A   W   M   R   P   N   D   L   I   W   N   Y  378

15282 TGG GTC AAC AAC TAC CTG CTC GGC AAC GAG CCG CCG GTG TTC GAC ATC CTG TTC TGG AAC  15341
  379 W   V   N   N   Y   L   L   G   N   E   P   P   V   F   D   I   L   F   W   N  398

15342 AAC GAC ACC ACG CGC CTG CCG GCC GCC TTC CAC GGC GAC CTG ATC GAA ATG TTC AAG AGC  15401
  399 N   D   T   T   R   L   P   A   A   F   H   G   D   L   I   E   M   F   K   S  418

15402 AAC CCG CTG ACC CGC CCG GAC GCC CTG GAG GTT TGC GGC ACT CCG ATC GAC CTG AAA CAG  15461
  419 N   P   L   T   R   P   D   A   L   E   V   C   G   T   P   I   D   L   K   Q  438

15462 GTC AAA TGC GAC ATC TAC AGC CTT GCC GGC ACC AAC GAC CAC ATC ACC CCG TGG CAG TCA  15521
  439 V   K   C   D   I   Y   S   L   A   G   T   N   D   H   I   T   P   W   Q   S  458
```

FIG. 23J

```
15522 TGC TAC CGC TCG GCG CAC CTG TTC GGC GGC AAG ATC GAG TTC GTG CTG TCC AAC AGC GGC 15581
  459 C   Y   R   S   A   H   L   F   G   G   K   I   E   F   V   L   S   N   S   G   478

15582 CAC ATC CAG AGC ATC CTC AAC CCG CCA GGC AAC CCC AAG GCG CGC TTC ATG ACC GGT GCC 15641
  479 H   I   Q   S   I   L   N   P   P   G   N   P   K   A   R   F   M   T   G   A   498

15642 GAT CGC CCG GGT GAC CCG GTG GCC TGG CAG GAA AAC GCC ACC AAG CAT GCC GAC TCC TGG 15701
  499 D   R   P   G   D   P   V   A   W   Q   E   N   A   T   K   H   A   D   S   W   518

15702 TGG CTG CAC TGG CAA AGC TGG CTG GGC GAG CGT GCC GGC GAG CTG GAA AAG GCG CCG ACC 15761
  519 W   L   H   W   Q   S   W   L   G   E   R   A   G   E   L   E   K   A   P   T   538

15762 CGC CTG GGC AAC CGT GCC TAT GCC GCT GGC GAG GCA TCC CCG GGC ACC TAC GTT CAC GAG 15821
  539 R   L   G   N   R   A   Y   A   A   G   E   A   S   P   G   T   Y   V   H   E   558

15822 CGT TGA GCTGCAGCGCCGTGGCCACCTGCGGGACGCCACGGTGTTGAATTC                          15872
  559 R   *                                                                        560
```

FIG. 23K

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90       100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

 BamHI
 ▼
GGATCCGACC GTGGGTGTGA ATCTCCGGGT GCTCGCCTCG TCCTGCCCCG TTACCTGTCC GCCTCCCGCT CCAGACCAGC GGGAGGCGGA CAGGGGCATG   100

SphI
▼
CCCGCCGGGC GGCTAACGGC CCGTGCGGCG TCCGTACGAC GAGCCTCGCG CGCCCTGGCG GCCCTTGGTC TGCCGGACCT GTGCGCGGGG TGCGCAGGGT   200

                                                                                          BstXI
                                                                                          ▼
TCGCCGCCGC GCGTGGGGCC GTATCTGCGG CTCCCGGGCA CGGCGGCCCT GCTCGTCTCC GAGTCATAGT CCCTGCCGCC GGCGCCACCG CCCTGGCCCG   300

     SphI
     ▼
GCATGCGCGT GCCGGGCGCC CCCGGCGCGT AACTCGGCTG GGAGGCCTGG AAAAGGGCGA TCCATTGGGT GAGCGTGAGG TCCTTCGGCA GTCCGCCGTC   400

   EcoRI
   ApoI
   ▼
CGGAATTCCG TGGCGGTCGG CGAGGGAACG GTAGGTCCGC TTGGGGATGT GGCGCCGGAG GATCTCCGCG AGGCCCCGTC CGGGGCCGGT GAAGACGGCT   500
```

FIG. 31A

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

TCGGCGAAGT TCTGGAAGGC GCGGCTCGCG CTCTCGGGCA GCAGGGGCTG GGGGCGTCGC CTGATCGTCA GGACGCCGCC GTCGACGCGG GGCATCGGAC   600

GGAACGACGA GGCGCGGACG CGGTCGTGGA CCGCGAACTC GTACCAGGGG GCCCAGGAGG TCGTGAGGAG CGATCCGCCG CTGCGACCGG CGCGTTTGCG   700

GGCGACCTCC CACTGCACTA TCAGGGCCGC CGACTGCCAG TTCGTCGATT CCAGGAGACT CCGGAGAATC TGGGTCGTGA TGCCGAAGGG AACGTTTCCG   800

                                                                              ApoI
                EcoRV           BsmI                                         BstBI
                  ▼               ▼                                         ▼  ▼
ACGACGGTGT CGATATCGCG CGGAATGCGG AAGTCGAGGA AATCACCCTG GAATACGGTG ACCCTCTCCC CTTCGAATTT CCGCCGCACA TGCGCGGCCC   900

AGTGCGGGTC CATCTCCACG ACCGTCACGG TGTCGAAGGA GCGCACCAAC TCCTCGGTTA TCGCGCCCTT TCCGGGGCCG ATTTCGAGAA CGTTCCTACC  1000
```

FIG. 31B

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GTCCCCCTCG ACATGCGTGA CGAGATTGCG CACGGCTCTG TCGTCCTGAA GGAAGTTCTG GCCTAATTCG CGGCGAAGGG TGTCGCGGTC CGCTCGCCTC  1100

GGTATGGAGT CGCGCATTGC CATgaacgat ccctccctg gatgccgtgg tcaatggact tggcacggac catacctcac ggtccgtcgg acgaccggag  1200

    XmnI
    ▼
aagaagttca cgcacgggcg ttccggagta cgggagttgt gaacggccgc gacgaagtcg gtcgcggctc ggcgggcggt gacgagcgag gtccggagga  1300

acgcgacgaa gcagccgaac cccaagtgag gtgcgacgga gtgacattgg gggcatacgg agggttgtcg tacggagcgc actcaacgag gctccaggag  1400

ggaggggttg aacccgccgc cgactggcct tcgccgcccg cgcggccgga gtatgtcatg tcgggggtga aatcaagcca ttccccccggg atcggctgtt  1500
```

FIG. 31C

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

acccatccct ttacctggcg tggatttccc aacccttggt atagagcggg agacgacgcg acaccatgga gaccacgcac accacgagcg ccacccccg  1600

gccatcccga caaggggggt ccggctcgcc tcccgacacc catggcctgg ggtacacgcc aggtataggg ggaacgtagg gggagcatag gggggtgcc  1700

ctggggttgg gtgaaagcgc ggcttccgga gacggagccg gATGTCTTCA GCCGGAATTA CCAGGACCGG TGCGAGAACA CCGGTGACAG GGCGTGGGGC  1800
                                              M  S  S  A  G  I  T  R  T  G  A  R  T  P  V  T  G  R  G  A

GGCAGCGTGG GACACGGGGG AAGTGCGGGT CCGACGGGGG TTGCCCCCTG CCGGCCCCGA TCATGCGGAG CACTCCTTCT CTCGTGCTCC TACCGGTGAT  1900
 A  A  W  D  T  G  E  V  R  V  R  R  G  L  P  P  A  G  P  D  H  A  E  H  S  F  S  R  A  P  T  G  D

                                                                                  EcoRI
                                                                                  ApoI
                  XmnI
                   ▼                                                               ▼
GTGCGCGCCG AATTGATTCG TGGAGAGATG TCGACAGTGT CCAAGAGTGA GTCCGAGGAA TTCGTGTCCG TGTCGAACGA CGCCGGTTCC GCGCACGGCA  2000
V  R  A  E  L  I  R  G  E  M  S  T  V  S  K  S  E  S  E  E  F  V  S  V  S  N  D  A  G  S  A  H  G  T
```

FIG. 31D

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CAGCGGAACC CGTCGCCGTC GTCGGCATCT CCTGCCGGGT GCCCGGCGCC CGGGACCCGA GAGAGTTCTG GGAACTCCTG GCGGCAGGCG GCCAGGCCGT  2100
  A  E  P  V  A  V  V  G  I  S  C  R  V  P  G  A  R  D  P  R  E  F  W  E  L  L  A  A  G  G  Q  A  V

CACCGACGTC CCCGCGGACC GCTGGAACGC CGGCGACTTC TACGACCCGG ACCGCTCCGC CCCCGGCCGC TCGAACAGCC GGTGGGGCGG GTTCATCGAG  2200
  T  D  V  P  A  D  R  W  N  A  G  D  F  Y  D  P  D  R  S  A  P  G  R  S  N  S  R  W  G  G  F  I  E

GACGTCGACC GGTTCGACGC CGCCTTCTTC GGCATCTCGC CCCGCGAGGC CGCGGAGATG GACCCGCAGC AGCGGCTCGC CCTGGAGCTG GGCTGGGAGG  2300
  D  V  D  R  F  D  A  A  F  F  G  I  S  P  R  E  A  A  E  M  D  P  Q  Q  R  L  A  L  E  L  G  W  E  A

CCCTGGAGCG CGCCGGGATC GACCCGTCCT CGCTCACCGG CACCCGCACC GGCGTCTTCG CCGGCGCCAT CTGGGACGAC TACGCCACCC TGAAGCACCG  2400
  L  E  R  A  G  I  D  P  S  S  L  T  G  T  R  T  G  V  F  A  G  A  I  W  D  D  Y  A  T  L  K  H  R

CCAGGGCGGC GCCGCGATCA CCCCGCACAC CGTCACCGGC CTCCACCGCG GCATCATCGC GAACCGACTC TCGTACACGC TCGGGCTCCG CGGCCCCAGC  2500
  Q  G  G  A  A  I  T  P  H  T  V  T  G  L  H  R  G  I  I  A  N  R  L  S  Y  T  L  G  L  R  G  P  S
```

FIG. 31E

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

ATGGTCGTCG ACTCCGGCCA GTCCTCGTCG CTCGTCGCCG TCCACCTCGC GTGCGAGAGC CTGCGGCGCG GCGAGTCCGA GCTCGCCCTC GCCGGCGGCG  2600
M  V  V  D   S  G  Q   S  S  S   L  V  A  V   H  L  A   C  E  S   L  R  R  G   E  S  E   L  A  L   A  G  G  V

TCTCGCTCAA CCTGGTGCCG GACAGCATCA TCGGGGCGAG CAAGTTCGGC GGCCTCTCCC CCGACGGCCG CGCCTACACC TTCGACGCGC GCGCCAACGG  2700
 S  L  N   L  V  P   D  S  I  I   G  A  S   K  F  G   G  L  S  P   D  G  R   A  Y  T   F  D  A  R   A  N  G

   SnaBI
   ▼
CTACGTACGC GGCGAGGGCG GCGGTTTCGT CGTCCTGAAG CGCCTCTCCC GGGCCGTCGC CGACGGCGAC CCGGTGCTCG CCGTGATCCG GGGCAGCGCC  2800
 Y  V  R   G  E  G  G   G  F  V   V  L  K   R  L  S  R   A  V  A   D  G  D   P  V  L  A   V  I  R   G  S  A

GTCAACAACG GCGGCGCCGC CCAGGGCATG ACGACCCCCG ACGCGCAGGC GCAGGAGGCC GTGCTCCGCG AGGCCCACGA GCGGGCCGGG ACCGCGCCGG  2900
V  N  N  G   G  A  A   Q  G  M   T  T  P  D   A  Q  A   Q  E  A   V  L  R  E   A  H  E   R  A  G   T  A  P  A

CCGACGTGCG GTACGTCGAG CTGCACGGCA CCGGCACCCC CGTGGGCGAC CCGATCGAGG CCGCTGCGCT CGGCGCCGCC CTCGGCACCG GCCGCCCGGC  3000
 D  V  R   Y  V  E   L  H  G  T   G  T  P   V  G  D   P  I  E  A   A  A  L   G  A  A   L  G  T  G   R  P  A
```

FIG. 31F

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CGGACAGCCG CTCCTGGTCG GCTCGGTCAA GACGAACATC GGCCACCTGG AGGGCGCGGC CGGCATCGCC GGCCTCATCA AGGCCGTCCT GGCGGTCCGC  3100
 G  Q  P   L  L  V  G   S  V  K   T  N  I   G  H  L  E   G  A  A   G  I  A   G  L  I  K   A  V  L   A  V  R

GGTCGCGCGC TGCCCGCCAG CCTGAACTAC GAGACCCCGA ACCCGGCGAT CCCGTTCGAG GAACTGAACC TCCGGGTGAA CACGGAGTAC CTGCCGTGGG  3200
G  R  A  L   P  A  S   L  N  Y   E  T  P  N   P  A  I   P  F  E   E  L  N  L   R  V  N   T  E  Y   L  P  W  E

AGCCGGAGCA CGACGGGCAG CGGATGGTCG TCGGCGTGTC CTCGTTCGGC ATGGGCGGCA CGAACGCGCA TGTCGTGCTC GAAGAGGCCC CCGGGGGTTG  3300
  P  E  H   D  G  Q   R  M  V  V   G  V  S   S  F  G   M  G  G  T   N  A  H   V  V  L   E  E  A  P   G  G  C

TCGAGGTGCT TCGGTCGTGG AGTCGACGGT CGGCGGGTCG GCGGTCGGCG GCGGTGTGGT GCCGTGGGTG GTGTCGGCGA AGTCCGCTGC CGCGCTGGAC  3400
 R  G  A   S  V  V  E   S  T  V   G  G  S   A  V  G  G   G  V  V   P  W  V   V  S  A  K   S  A  A   A  L  D

GCGCAGATCG AGCGGCTTGC CGCGTTCGCC TCGCGGGATC GTACGGATGG TGTCGACGCG GGCGCTGTCG ATGCGGGTGC TGTCGATGCG GGTGCTGTCG  3500
A  Q  I  E   R  L  A   A  F  A   S  R  D  R   T  D  G   V  D  A   G  A  V  D   A  G  A   V  D  A   G  A  V  A
```

FIG. 31G

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                                                  Eco47III   Bsu361
                                                                                     ▼          ▼
CTCGCGTACT GGCCGGCGGG CGTGCTCAGT TCGAGCACCG GGCCGTCGTC GTCGGCAGCG GGCCGGACGA TCTGGCGGCA GCGCTGGCCG CGCCTGAGGG  3600
  R  V  L   A  G  G   R  A  Q  F   E  H  R   A  V  V   V  G  S  G   P  D  D   L  A  A   A  L  A  A   P  E  G

TCTGGTCCGG GGCGTGGCTT CCGGTGTCGG GCGAGTGGCG TTCGTGTTCC CCGGGCAGGG CACGCAGTGG GCCGGCATGG GTGCCGAACT GCTGGACTCT  3700
  L  V  R   G  V  A  S   G  V  G   R  V  A   F  V  F  P   G  Q  G   T  Q  W   A  G  M  G   A  E  L   L  D  S

                                      BsmI
                                       ▼
TCCGCGGTGT TCGCGGCGGC CATGGCCGAA TGCGAGGCCG CACTCTCCCC GTACGTCGAC TGGTCGCTGG AGGCCGTCGT ACGGCAGGCC CCCGGTGCGC  3800
  S  A  V  F   A  A  A   M  A  E   C  E  A  A   L  S  P   Y  V  D   W  S  L  E   A  V  V   R  Q  A   P  G  A  P

CCACGCTGGA GCGGGTCGAT GTCGTGCAGC CTGTGACGTT CGCCGTCATG GTCTCGCTGG CTCGCGTGTG GCAGCACCAC GGGGTGACGÇ CCCAGGCGGT  3900
  T  L  E   R  V  D   V  V  Q  P   V  T  F   A  V  M   V  S  L  A   R  V  W   Q  H  H   G  V  T  P   Q  A  V

CGTCGGCCAC TCGCAGGGCG AGATCGCCGC CGCGTACGTC GCCGGTGCCC TGAGCCTGGA CGACGCCGCT CGTGTCGTGA CCCTGCGCAG CAAGTCCATC  4000
  V  G  H   S  Q  G  E   I  A  A   A  Y  V   A  G  A  L   S  L  D   D  A  A   R  V  V  T   L  R  S   K  S  I
```

FIG. 31H

pikPKS Sequence

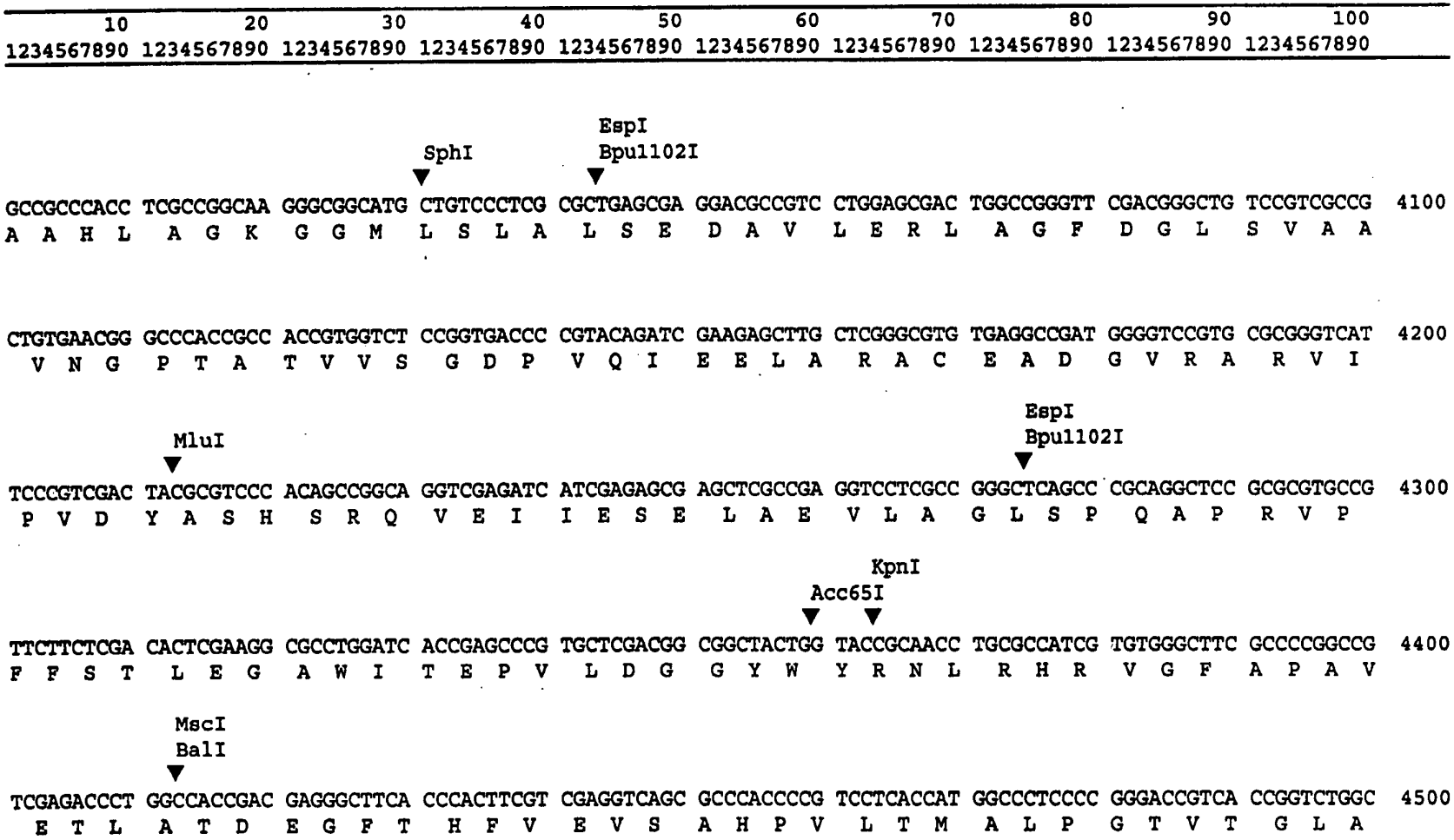

FIG. 311

pikPKS Sequence

```
          10        20        30        40        50        60        70        80        90       100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                      MaeI
                                      BfaI
                                       ▼
GACCCTGCGT CGCGACAACG GCGGTCAGGA CCGCCTAGTC GCCTCCCTCG CCGAAGCATG GGCCAACGGA CTCGCGGTCG ACTGGAGCCC GCTCCTCCCC  4600
 T  L  R   R  D  N  G   G  Q  D   R  L  V   A  S  L  A   E  A  W   A  N  G   L  A  V  D   W  S  P   L  L  P

                                      MluI
                                       ▼
TCCGCGACCG GCCACCACTC CGACCTCCCC ACCTACGCGT TCCAGACCGA GCGCCACTGG CTGGGCGAGA TCGAGGCGCT CGCCCCGGCG GGCGAGCCGG  4700
 S  A  T  G   H  H  S   D  L  P   T  Y  A  F   Q  T  E   R  H  W   L  G  E  I   E  A  L   A  P  A   G  E  P  A

CGGTGCAGCC CGCCGTCCTC CGCACGGAGG CGGCCGAGCC GGCGGAGCTC GACCGGGACG AGCAGCTGCG CGTGATCCTG GACAAGGTCC GGGCGCAGAC  4800
  V  Q  P   A  V  L   R  T  E  A   A  E  P   A  E  L   D  R  D  E   Q  L  R   V  I  L   D  K  V  R   A  Q  T

GGCCCAGGTG CTGGGGTACG CGACAGGCGG GCAGATCGAG GTCGACCGGA CCTTCCGTGA GGCCGGTTGC ACCTCCCTGA CCGGCGTGGA CCTGCGCAAC  4900
 A  Q  V   L  G  Y  A   T  G  G   Q  I  E   V  D  R  T   F  R  E   A  G  C   T  S  L  T   G  V  D   L  R  N

                                                                                                    ApaLI
                                                                                                     ▼
CGGATCAACG CCGCCTTCGG CGTACGGATG GCGCCGTCCA TGATCTTCGA CTTCCCCACC CCCGAGGCTC TCGCGGAGCA GCTGCTCCTC GTCGTGCACG  5000
R  I  N  A   A  F  G   V  R  M   A  P  S  M   I  F  D   F  P  T   P  E  A  L   A  E  Q   L  L  L   V  V  H  G
```

FIG. 31J

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GGGAGGCGGC GGCGAACCCG GCCGGTGCGG AGCCGGCTCC GGTGGCGGCG GCCGGTGCCG TCGACGAGCC GGTGGCGATC GTCGGCATGG CCTGCCGCCT  5100
  E  A  A  A  N  P  A  G  A  E  P  A  P  V  A  A  A  G  A  V  D  E  P  V  A  I  V  G  M  A  C  R  L

GCCCGGTGGG GTCGCCTCGC CGGAGGACCT GTGGCGGCTG GTGGCCGGCG GCGGGGACGC GATCTCGGAG TTCCCGCAGG ACCGCGGCTG GGACGTGGAG  5200
  P  G  G  V  A  S  P  E  D  L  W  R  L  V  A  G  G  G  D  A  I  S  E  F  P  Q  D  R  G  W  D  V  E

                BamHI
                ▼
GGGCTGTACC ACCCGGATCC GGAGCACCCC GGCACGTCGT ACGTCCGCCA GGGCGGTTTC ATCGAGAACG TCGCCGGCTT CGACGCGGCC TTCTTCGGGA  5300
  G  L  Y  H  P  D  P  E  H  P  G  T  S  Y  V  R  Q  G  G  F  I  E  N  V  A  G  F  D  A  A  F  F  G  I

TCTCGCCGCG CGAGGCCCTC GCCATGGACC CGCAGCAGCG GCTCCTCCTC GAAACCTCCT GGGAGGCCGT CGAGGACGCC GGGATCGACC CGACCTCCCT  5400
  S  P  R  E  A  L  A  M  D  P  Q  Q  R  L  L  L  E  T  S  W  E  A  V  E  D  A  G  I  D  P  T  S  L

GCGGGGACGG CAGGTCGGCG TCTTCACTGG GGCGATGACC CACGAGTACG GGCCGAGCCT GCGGGACGGC GGGGAAGGCC TCGACGGCTA CCTGCTGACC  5500
  R  G  R  Q  V  G  V  F  T  G  A  M  T  H  E  Y  G  P  S  L  R  D  G  G  E  G  L  D  G  Y  L  L  T
```

FIG. 31K

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GGCAACACGG CCAGCGTGAT GTCGGGCCGC GTCTCGTACA CACTCGGCCT TGAGGGCCCC GCCCTGACGG TGGACACGGC CTGCTCGTCG TCGCTGGTCG  5600
G  N  T  A  S  V  M  S  G  R  V  S  Y  T  L  G  L  E  G  P  A  L  T  V  D  T  A  C  S  S  S  L  V  A

CCCTGCACCT CGCCGTGCAG GCCCTGCGCA AGGGCGAGGT CGACATGGCG CTCGCCGGCG GCGTGGCCGT GATGCCCACG CCCGGGATGT TCGTCGAGTT  5700
  L  H  L  A  V  Q  A  L  R  K  G  E  V  D  M  A  L  A  G  G  V  A  V  M  P  T  P  G  M  F  V  E  F

                                                  XmnI
                                                  ▼
CAGCCGGCAG CGCGGGCTGG CCGGGGACGG CCGGTCGAAG GCGTTCGCCG CGTCGGCGGA CGGCACCAGC TGGTCCGAGG GCGTCGGCGT CCTCCTCGTC  5800
 S  R  Q  R  G  L  A  G  D  G  R  S  K  A  F  A  A  S  A  D  G  T  S  W  S  E  G  V  G  V  L  L  V

GAGCGCCTGT CGGACGCCCG CCGCAACGGA CACCAGGTCC TCGCGGTCGT CCGCGGCAGC GCCTTGAACC AGGACGGCGC GAGCAACGGC CTCACGGCTC  5900
E  R  L  S  D  A  R  R  N  G  H  Q  V  L  A  V  V  R  G  S  A  L  N  Q  D  G  A  S  N  G  L  T  A  P

CGAACGGGCC CTCGCAGCAG CGCGTCATCC GGCGCGCGCT GGCGGACGCC CGGCTGACGA CCTCCGACGT GGACGTCGTC GAGGCACACG GCACGGGCAC  6000
  N  G  P  S  Q  Q  R  V  I  R  R  A  L  A  D  A  R  L  T  T  S  D  V  D  V  V  E  A  H  G  T  G  T
```

FIG. 31L

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GCGACTCGGC GACCCGATCG AGGCGCAGGC CCTGATCGCC ACCTACGGCC AGGGCCGTGA CGACGAACAG CCGCTGCGCC TCGGGTCGTT GAAGTCCAAC  6100
 R  L  G   D  P  I  E   A  Q  A   L  I  A   T  Y  G  Q   G  R  D   D  E  Q   P  L  R  L   G  S  L   K  S  N

ATCGGGCACA CCCAGGCCGC GGCCGGCGTC TCCGGTGTCA TCAAGATGGT CCAGGCGATG CGCCACGGAC TGCTGCCGAA GACGCTGCAC GTCGACGAGC  6200
 I  G  H  T   Q  A  A   A  G  V   S  G  V  I   K  M  V   Q  A  M   R  H  G  L   L  P  K   T  L  H   V  D  E  P

CCTCGGACCA GATCGACTGG TCGGCTGGCG CCGTGGAACT CCTCACCGAG GCCGTCGACT GGCCGGAGAA GCAGGACGGC GGGCTGCGCC GGGCCGCCGT  6300
 S  D  Q   I  D  W   S  A  G  A   V  E  L   L  T  E   A  V  D  W   P  E  K   Q  D  G   G  L  R  R   A  A  V

CTCCTCCTTC GGGATCAGCG GCACCAATGC GCATGTGGTG CTCGAAGAGG CCCCGGTGGT TGTCGAGGGT GCTTCGGTCG TCGAGCCGTC GGTTGGCGGG  6400
 S  S  F   G  I  S  G   T  N  A   H  V  V   L  E  E  A   P  V  V   V  E  G   A  S  V  V   E  P  S   V  G  G

                                                                                          BsmI
                                                                                           ▼
TCGGCGGTCG GCGGCGGTGT GACGCCTTGG GTGGTGTCGG CGAAGTCCGC TGCCGCGCTC GACGCGCAGA TCGAGCGGCT TGCCGCATTC GCCTCGCGGG  6500
 S  A  V  G   G  G  V   T  P  W   V  V  S  A   K  S  A   A  A  L   D  A  Q  I   E  R  L   A  A  F   A  S  R  D
```

FIG. 31M

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

ATCGTACGGA TGACGCCGAC GCCGGTGCTG TCGACGCGGG CGCTGTCGCT CACGTACTGG CTGACGGGCG TGCTCAGTTC GAGCACCGGG CCGTCGCGCT  6600
  R  T  D   D  A  D   A  G  A  V   D  A  G   A  V  A   H  V  L  A   D  G  R   A  Q  F   E  H  R  A   V  A  L

                                                                    XmnI
                                                                     ▼
CGGCGCCGGG GCGGACGACC TCGTACAGGC GCTGGCCGAT CCGGACGGGC TGATACGCGG AACGGCTTCC GGTGTCGGGC GAGTGGCGTT CGTGTTCCCC  6700
  G  A  G   A  D  D  L   V  Q  A   L  A  D   P  D  G  L   I  R  G   T  A  S   G  V  G  R   V  A  F   V  F  P

GGTCAGGGCA CGCAGTGGGC TGGCATGGGT GCCGAACTGC TGGACTCTTC CGCGGTGTTC GCGGCGGCCA TGGCCGAGTG TGAGGCCGCG CTGTCCCCGT  6800
  G  Q  G  T   Q  W  A   G  M  G   A  E  L  L   D  S  S   A  V  F   A  A  A  M   A  E  C   E  A  A   L  S  P  Y

ACGTCGACTG GTCGCTGGAG GCCGTCGTAC GGCAGGCCCC CGGTGCGCCC ACGCTGGAGC GGGTCGATGT CGTGCAGCCT GTGACGTTCG CCGTCATGGT  6900
  V  D  W   S  L  E   A  V  V  R   Q  A  P   G  A  P   T  L  E  R   V  D  V   V  Q  P   V  T  F  A   V  M  V

CTCGCTGGCT CGCGTGTGGC AGCACCACGG TGTGACGCCC CAGGCGGTCG TCGGCCACTC GCAGGGCGAG ATCGCCGCCG CGTACGTCGC CGGAGCCCTG  7000
  S  L  A   R  V  W  Q   H  H  G   V  T  P   Q  A  V  V   G  H  S   Q  G  E   I  A  A  A   Y  V  A   G  A  L
```

FIG. 31N

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90       100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                                                  SphI
                                                                                  ▼
CCCCTGGACG ACGCCGCCCG CGTCGTCACC CTGCGCAGCA AGTCCATCGC CGCCCACCTC GCCGGCAAGG GCGGCATGCT GTCCCTCGCG CTGAACGAGG  7100
P  L  D  D   A  A  R   V  V  T   L  R  S  K   S  I  A   A  H  L   A  G  K  G   G  M  L   S  L  A   L  N  E  D

ACGCCGTCCT GGAGCGACTG AGTGACTTCG ACGGGCTGTC CGTCGCCGCC GTCAACGGGC CCACCGCCAC TGTCGTGTCG GGTGACCCCG TACAGATCGA  7200
  A  V  L   E  R  L   S  D  F  D   G  L  S   V  A  A   V  N  G  P   T  A  T   V  V  S   G  D  P  V   Q  I  E

                                                                  MluI
                                                                  ▼
AGAGCTTGCT CAGGCGTGCA AGGCGGACGG ATTCCGCGCG CGGATCATTC CCGTCGACTA CGCGTCCCAC AGCCGGCAGG TCGAGATCAT CGAGAGCGAG  7300
 E  L  A   Q  A  C  K   A  D  G   F  R  A   R  I  I  P   V  D  Y   A  S  H   S  R  Q  V   E  I  I   E  S  E

CTCGCCCAGG TCCTCGCCGG TCTCAGCCCG CAGGCCCCGC GCGTGCCGTT CTTCTCGACG CTCGAAGGCA CCTGGATCAC CGAGCCCGTC CTCGACGGCA  7400
L  A  Q  V   L  A  G   L  S  P   Q  A  P  R   V  P  F   F  S  T   L  E  G  T   W  I  T   E  P  V   L  D  G  T

          KpnI
Acc65I    ▼
▼         ▼
CCTACTGGTA CCGCAACCTC CGTCACCGCG TCGGCTTCGC CCCCGCCATC GAGACCCTGG CCGTCGACGA GGGCTTCACG CACTTCGTCG AGGTCAGCGC  7500
 Y  W  Y   R  N  L   R  H  R  V   G  F  A   P  A  I   E  T  L   A  V  D  E   G  F  T   H  F  V  E   V  S  A
```

FIG. 310

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CCACCCCGTC CTCACCATGA CCCTCCCCGA GACCGTCACC GGCCTCGGCA CCCTCCGTCG CGAACAGGGA GGCCAAGAGC GTCTGGTCAC CTCGCTCGCC  7600
 H  P  V   L  T  M  T   L  P  E   T  V  T   G  L  G  T   L  R  R   E  Q  G   G  Q  E  R   L  V  T   S  L  A

                                                                                                  Eco47III
                                                                                                      ▼
GAGGCGTGGG TCAACGGGCT TCCCGTGGCA TGGACTTCGC TCCTGCCCGC CACGGCCTCC CGCCCCGGTC TGCCCACCTA CGCCTTCCAG GCCGAGCGCT  7700
E  A  W  V   N  G  L   P  V  A   W  T  S  L   L  P  A   T  A  S   R  P  G  L   P  T  Y   A  F  Q   A  E  R  Y

      XhoI                          MscI
      PaeR7I                        BalI
        ▼                             ▼
ACTGGCTCGA GAACACTCCC GCCGCCCTGG CCACCGGCGA CGACTGGCGC TACCGCATCG ACTGGAAGCG CCTCCCGGCC GCCGAGGGGT CCGAGCGCAC  7800
  W  L  E   N  T  P   A  A  L  A   T  G  D   D  W  R   Y  R  I  D   W  K  R   L  P  A   A  E  G  S   E  R  T

CGGCCTGTCC GGCCGCTGGC TCGCCGTCAC GCCGGAGGAC CACTCCGCGC AGGCCGCCGC CGTGCTCACC GCGCTGGTCG ACGCCGGGGC GAAGGTCGAG  7900
 G  L  S   G  R  W  L   A  V  T   P  E  D   H  S  A  Q   A  A  A   V  L  T   A  L  V  D   A  G  A   K  V  E

GTGCTGACGG CCGGGGCGGA CGACGACCGT GAGGCCCTCG CCGCCCGGCT CACCGCACTG ACGACCGGTG ACGGCTTCAC CGGCGTGGTC TCGCTCCTCG  8000
V  L  T  A   G  A  D   D  D  R   E  A  L  A   A  R  L   T  A  L   T  T  G  D   G  F  T   G  V  V   S  L  L  D
```

FIG. 31P

pikPKS Sequence

```
      10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

ACGGACTCGT ACCGCAGGTC GCCTGGGTCC AGGCGCTCGG CGACGCCGGA ATCAAGGCGC CCCTGTGGTC CGTCACCCAG GGCGCGGTCT CCGTCGGACG  8100
  G  L  V   P  Q  V   A  W  V  Q   A  L  G   D  A  G   I  K  A  P   L  W  S   V  T  Q   G  A  V  S   V  G  R

TCTCGACACC CCCGCCGACC CCGACCGGGC CATGCTCTGG GGCCTCGGCC GCGTCGTCGC CCTTGAGCAC CCCGAACGCT GGGCCGGCCT CGTCGACCTC  8200
  L  D  T   P  A  D  P   D  R  A   M  L  W   G  L  G  R   V  V  A   L  E  H   P  E  R  W   A  G  L   V  D  L

                                                                                    BsaBI
                                                                                      ▼
CCCGCCCAGC CCGATGCCGC CGCCCTCGCC CACCTCGTCA CCGCACTCTC CGGCGCCACC GGCGAGGACC AGATCGCCAT CCGCACCACC GGACTCCACG  8300
  P  A  Q   P  D  A   A  A  L   A  H  L  V   T  A  L   S  G  A   T  G  E   D  Q  I   A  I  R   T  T  G   L  H  A

CCCGCCGCCT CGCCCGCGCA CCCCTCCACG GACGTCGGCC CACCCGCGAC TGGCAGCCCC ACGGCACCGT CCTCATCACC GGCGGCACCG GAGCCCTCGG  8400
  R  R  L   A  R  A   P  L  H  G   R  R  P   T  R  D   W  Q  P  H   G  T  V   L  I  T   G  G  T  G   A  L  G

CAGCCACGCC GCACGCTGGA TGGCCCACCA CGGAGCCGAA CACCTCCTCC TCGTCAGCCG CAGCGGCGAA CAAGCCCCCG GAGCCACCCA ACTCACCGCC  8500
  S  H  A   A  R  W  M   A  H  H   G  A  E   H  L  L  L   V  S  R   S  G  E   Q  A  P  G   A  T  Q   L  T  A
```

FIG. 31Q

pikPKS Sequence

```
          10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GAACTCACCG CATCGGGCGC CCGCGTCACC ATCGCCGCCT GCGACGTCGC CGACCCCCAC GCCATGCGCA CCCTCCTCGA CGCCATCCCC GCCGAGACGC  8600
E  L  T  A  S  G  A  R  V  T  I  A  A  C  D  V  A  D  P  H  A  M  R  T  L  L  D  A  I  P  A  E  T  P

CCCTCACCGC CGTCGTCCAC ACCGCCGGCG CGCTCGACGA CGGCATCGTG GACACGCTGA CCGCCGAGCA GGTCCGGCGG GCCCACCGTG CGAAGGCCGT  8700
  L  T  A  V  V  H  T  A  G  A  L  D  D  G  I  V  D  T  L  T  A  E  Q  V  R  R  A  H  R  A  K  A  V

                                                  MluI
                                                   ▼
CGGCGCCTCG GTGCTCGACG AGCTGACCCG GGACCTCGAC CTCGACGCGT TCGTGCTCTT CTCGTCCGTG TCGAGCACTC TGGGCATCCC CGGTCAGGGC  8800
  G  A  S  V  L  D  E  L  T  R  D  L  D  L  D  A  F  V  L  F  S  S  V  S  S  T  L  G  I  P  G  Q  G

AACTACGCCC CGCACAACGC CTACCTCGAC GCCCTCGCGG CTCGCCGCCG GGCCACCGGC CGGTCCGCCG TCTCGGTGGC CTGGGGACCG TGGGACGGTG  8900
N  Y  A  P  H  N  A  Y  L  D  A  L  A  A  R  R  R  A  T  G  R  S  A  V  S  V  A  W  G  P  W  D  G  G

GCGGCATGGC CGCCGGTGAC GGCGTGGCCG AGCGGCTGCG CAACCACGGC GTGCCCGGCA TGGACCCGGA ACTCGCCCTG GCCGCACTGG AGTCCGCGCT  9000
  G  M  A  A  G  D  G  V  A  E  R  L  R  N  H  G  V  P  G  M  D  P  E  L  A  L  A  A  L  E  S  A  L
```

FIG. 31R

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CGGCCGGGAC GAGACCGCGA TCACCGTCGC GGACATCGAC TGGGACCGCT TCTACCTCGC GTACTCCTCC GGTCGCCCGC AGCCCCTCGT CGAGGAGCTG  9100
 G  R  D  E  T  A  I  T  V  A  D  I  D  W  D  R  F  Y  L  A  Y  S  S  G  R  P  Q  P  L  V  E  E  L

                                                                                  BstXI
                                                                                    ▼
CCCGAGGTGC GGCGCATCAT CGACGCACGG GACAGCGCCA CGTCCGGACA GGGCGGGAGC TCCGCCCAGG GCGCCAACCC CCTGGCCGAG CGGCTGGCCG  9200
 P  E  V  R  R  I  I  D  A  R  D  S  A  T  S  G  Q  G  G  S  S  A  Q  G  A  N  P  L  A  E  R  L  A  A

CCGCGGCTCC CGGCGAGCGT ACGGAGATCC TCCTCGGTCT CGTACGGGCG CAGGCCGCCG CCGTGCTCCG GATGCGTTCG CCGGAGGACG TCGCCGCCGA  9300
 A  A  P  G  E  R  T  E  I  L  L  G  L  V  R  A  Q  A  A  A  V  L  R  M  R  S  P  E  D  V  A  A  D

CCGCGCCTTC AAGGACATCG GCTTCGACTC GCTCGCCGGT GTCGAGCTGC GCAACAGGCT GACCCGGGCG ACCGGGCTCC AGCTGCCCGC GACGCTCGTC  9400
 R  A  F  K  D  I  G  F  D  S  L  A  G  V  E  L  R  N  R  L  T  R  A  T  G  L  Q  L  P  A  T  L  V

TTCGACCACC CGACGCCGCT GGCCCTCGTG TCGCTGCTCC GCAGCGAGTT CCTCGGTGAC GAGGAGACGG CGGACGCCCG GCGGTCCGCG GCGCTGCCCG  9500
 F  D  H  P  T  P  L  A  L  V  S  L  L  R  S  E  F  L  G  D  E  E  T  A  D  A  R  R  S  A  A  L  P  A
```

FIG. 31S

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                                  BsaBI
                                                                    ▼
CGACTGTCGG TGCCGGTGCC GGCGCCGGCG CCGGCACCGA TGCCGACGAC GATCCGATCG CGATCGTCGC GATGAGCTGC CGCTACCCCG GTGACATCCG  9600
  T  V  G   A  G  A   G  A  G  A   G  T  D   A  D  D   D  P  I  A   I  V  A   M  S  C   R  Y  P  G   D  I  R

CAGCCCGGAG GACCTGTGGC GGATGCTGTC CGAGGGCGGC GAGGGCATCA CGCCGTTCCC CACCGACCGC GGCTGGGACC TCGACGGCCT GTACGACGCC  9700
  S  P  E   D  L  W   R  M  L  S   E  G  G   E  G  I   T  P  F   P  T  D   R  G  W   D  L  D   G  L  Y  D  A

GACCCGGACG CGCTCGGCAG GGCGTACGTC CGCGAGGGCG GGTTCCTGCA CGACGCGGCC GAGTTCGACG CGGAGTTCTT CGGCGTCTCG CCGCGCGAGG  9800
D  P  D  A   L  G  R   A  Y  V   R  E  G  G   F  L  H   D  A  A   E  F  D  A   E  F  F   G  V  S   P  R  E  A

     MscI
     BalI
      ▼
CGCTGGCCAT GGACCCGCAG CAGCGGATGC TCCTGACGAC GTCCTGGGAG GCCTTCGAGC GGGCCGGCAT CGAGCCGGCA TCGCTGCGCG GCAGCAGCAC  9900
  L  A  M   D  P  Q   Q  R  M  L   L  T  T   S  W  E   A  F  E  R   A  G  I   E  P  A   S  L  R  G   S  S  T

CGGTGTCTTC ATCGGCCTCT CCTACCAGGA CTACGCGGCC CGCGTCCCGA ACGCCCCGCG TGGCGTGGAG GGTTACCTGC TGACCGGCAG CACGCCGAGC 10000
  G  V  F   I  G  L  S   Y  Q  D   Y  A  A   R  V  P  N   A  P  R   G  V  E   G  Y  L  L   T  G  S   T  P  S
```

FIG. 31T

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GTCGCGTCGG GCCGTATCGC GTACACCTTC GGTCTCGAAG GGCCCGCGAC GACCGTCGAC ACCGCCTGCT CGTCGTCGCT GACCGCCCTG CACCTGGCGG 10100
V  A  S  G  R  I  A  Y  T  F  G  L  E  G  P  A  T  T  V  D  T  A  C  S  S  S  L  T  A  L  H  L  A  V

                        ApaLI
                          ▼
TGCGGGCGCT GCGCAGCGGC GAGTGCACGA TGGCGCTCGC CGGTGGCGTG GCGATGATGG CGACCCCGCA CATGTTCGTG GAGTTCAGCC GTCAGCGGGC 10200
  R  A  L  R  S  G  E  C  T  M  A  L  A  G  G  V  A  M  M  A  T  P  H  M  F  V  E  F  S  R  Q  R  A

GCTCGCCCCG GACGGCCGCA GCAAGGCCTT CTCGGCGGAC GCCGACGGGT TCGGCGCCGC GGAGGGCGTC GGCCTGCTGC TCGTGGAGCG GCTCTCGGAC 10300
  L  A  P  D  G  R  S  K  A  F  S  A  D  A  D  G  F  G  A  A  E  G  V  G  L  L  L  V  E  R  L  S  D

                                                  KpnI
                                                Acc65I
                                                  ▼  ▼
GCGCGGCGCA ACGGTCACCC GGTGCTCGCC GTGGTCCGCG GTACCGCCGT CAACCAGGAC GGCGCCAGCA ACGGGCTGAC CGCGCCCAAC GGACCCTCGC 10400
A  R  R  N  G  H  P  V  L  A  V  V  R  G  T  A  V  N  Q  D  G  A  S  N  G  L  T  A  P  N  G  P  S  Q

AGCAGCGGGT GATCCGGCAG GCGCTCGCCG ACGCCCGGCT GGCACCCGGC GACATCGACG CCGTCGAGAC GCACGGCACG GGAACCTCGC TGGGCGACCC 10500
  Q  R  V  I  R  Q  A  L  A  D  A  R  L  A  P  G  D  I  D  A  V  E  T  H  G  T  G  T  S  L  G  D  P
```

FIG. 31U

pikPKS Sequence

```
         20        30        40        50        60        70        80        90       100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CAGGGCCTCC AGGCCACGTA CGGCAAGGAG CGGCCCGCGG AACGGCCGCT CGCCATCGGC TCCGTGAAGT CCAACATCGG ACACACCCAG 10600
 Q  G  L  Q   A  T  Y   G  K  E   R  P  A  E   R  P  L   A  I  G   S  V  K  S   N  I  G   H  T  Q

GTGCGGCGGG CATCATCAAG ATGGTCCTCG CGATGCGCCA CGGCACCCTG CCGAAGACCC TCCACGCCGA CGAGCCGAGC CCGCACGTCG 10700
 A  A  G   I  I  K   M  V  L  A   M  R  H   G  T  L   P  K  T  L   H  A  D   E  P  S   P  H  V  D

CAGCGGCCTG GCCCTCGTCA CCGAGCCGAT CGACTGGCCG GCCGGCACCG GTCCGCGCCG CGCCGCCGTC TCCTCCTTCG GCATCAGCGG 10800
 S  G  L   A  L  V  T   E  P  I   D  W  P   A  G  T  G   P  R  R   A  A  V   S  S  F  G   I  S  G

                                                                    Bsu36I
                                                                      ▼
CACGTCGTGC TGGAGCAGGC GCCGGATGCT GCTGGTGAGG TGCTTGGGGC CGATGAGGTG CCTGAGGTGT CTGAGACGGT AGCGATGGCT 10900
 H  V  V  L   E  Q  A   P  D  A   A  G  E  V   L  G  A   D  E  V   P  E  V  S   E  T  V   A  M  A

GGACCTCCGA GGTCGCTGAG GGCTCTGAGG CCTCCGAGGC CCCCGCGGCC CCCGGCAGCC GTGAGGCGTC CTCCCCGGG CACCTGCCCT 11000
 T  S  E   V  A  E   G  S  E  A   S  E  A   P  A  A   P  G  S  R   E  A  S   L  P  G   H  L  P  W
```

FIG. 31V

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CCGACGTGGC CCTGCCGGCC GAGGTCGCCG GTGCCGAGGG CGCGCGGTTC GGCCTTCACC CGGCGCTGCT CGACGCCGCC GTGCAGGCGG CCGGTGCGGG 23600
  D  V  A  L  P  A  E  V  A  G  A  E  G  A  R  F  G  L  H  P  A  L  L  D  A  A  V  Q  A  A  G  A  G

CCGGGGCGTT CGGCGCGGGC ACGCGGCTGC CGTTCGCCTG GAGCGGGATC TCCTGTACGC GGTCGGCGCC ACCGCCCTCC GCGTGCGGCT GGCCCCCGCC 23700
  R  G  V  R  R  G  H  A  A  A  V  R  L  E  R  D  L  L  Y  A  V  G  A  T  A  L  R  V  R  L  A  P  A

GGCCCGGACA CGGTGTCCGT GAGCGCCGCC GACTCCTCCG GGCAGCCGGT GTTCGCCGCG GACTCCCTCA CGGTGCTGCC CGTCGACCCC GCGCAGCTGG 23800
G  P  D  T  V  S  V  S  A  A  D  S  S  G  Q  P  V  F  A  A  D  S  L  T  V  L  P  V  D  P  A  Q  L  A

CGGCCTTCAG CGACCCGACT CTGGACGCGC TGCACCTGCT GGAGTGGACC GCCTGGGACG GTGCCGCGCA GGCCCTGCCC GGCGCGGTCG TGCTGGGCGG 23900
  A  F  S  D  P  T  L  D  A  L  H  L  L  E  W  T  A  W  D  G  A  A  Q  A  L  P  G  A  V  V  L  G  G

CGACGCCGAC GGTCTCGCCG CGGCGCTGCG CGCCGGTGGC ACCGAGGTCC TGTCCTTCCC GGACCTTACG GACCTGGTGG AGGCCGTCGA CCGGGGCGAG 24000
  D  A  D  G  L  A  A  A  L  R  A  G  G  T  E  V  L  S  F  P  D  L  T  D  L  V  E  A  V  D  R  G  E
```

FIG. 31W

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                        SphI
                                        ▼
GCGCTCTTCC GGCTCGTCGA GAGCTGGGGC ATGCGGCCGG CCGCACTGCT CGGTCACTCG GTCGGCGAGA TCGCCGCCGC GCACGTCGCC GGTGTGTTCT 11600
A  L  F  R  L  V  E  S  W  G  M  R  P  A  A  L  L  G  H  S  V  G  E  I  A  A  A  H  V  A  G  V  F  S

CGCTCGCCGA CGCCGCCCGC CTGGTCGCCG CGCGCGGCCG GCTCATGCAG GAGCTGCCCG CCGGTGGCGC GATGCTCGCC GTCCAGGCCG CGGAGGACGA 11700
  L  A  D  A  A  R  L  V  A  A  R  G  R  L  M  Q  E  L  P  A  G  G  A  M  L  A  V  Q  A  A  E  D  E

GATCCGCGTG TGGCTGGAGA CGGAGGAGCG GTACGCGGGA CGTCTGGACG TCGCCGCCGT CAACGGCCCC GAGGCCGCCG TCCTGTCCGG CGACGCGGAC 11800
 I  R  V  W  L  E  T  E  E  R  Y  A  G  R  L  D  V  A  A  V  N  G  P  E  A  A  V  L  S  G  D  A  D

                                                                                                  SphI
                                                                                                  ▼
GCGGCGCGGG AGGCGGAGGC GTACTGGTCC GGGCTCGGCC GCAGGACCCG CGCGCTGCGG GTCAGCCACG CCTTCCACTC CGCGCACATG GACGGCATGC 11900
A  A  R  E  A  E  A  Y  W  S  G  L  G  R  R  T  R  A  L  R  V  S  H  A  F  H  S  A  H  M  D  G  M  L

TCGACGGGTT CCGCGCCGTC CTGGAGACGG TGGAGTTCCG GCGCCCCTCC CTGACCGTGG TCTCGAACGT CACCGGCCTG GCCGCCGGCC CGGACGACCT 12000
  D  G  F  R  A  V  L  E  T  V  E  F  R  R  P  S  L  T  V  V  S  N  V  T  G  L  A  A  G  P  D  D  L
```

FIG. 31X

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                ScaI
                ▼
GTGCGACCCC GAGTACTGGG TCCGGCACGT CCGCGGCACC GTCCGCTTCC TCGACGGCGT CCGTGTCCTG CGCGACCTCG GCGTGCGGAC CTGCCTGGAG 12100
 C  D  P   E  Y  W  V   R  H  V   R  G  T   V  R  F  L   D  G  V   R  V  L   R  D  L  G   V  R  T   C  L  E

CTGGGCCCCG ACGGGGTCCT CACCGCCATG GCGGCCGACG GCCTCGCGGA CACCCCCGCG GATTCCGCTG CCGGCTCCCC CGTCGGCTCT CCCGCCGGCT 12200
L  G  P  D   G  V  L   T  A  M   A  A  D  G   L  A  D   T  P  A   D  S  A  A   G  S  P   V  G  S   P  A  G  S

CTCCCGCCGA CTCCGCCGCC GGCGCGCTCC GGCCCCGGCC GCTGCTCGTG GCGCTGCTGC GCCGCAAGCG GTCGGAGACC GAGACCGTCG CGGACGCCCT 12300
  P  A  D   S  A  A   G  A  L  R   P  R  P   L  L  V   A  L  L  R   R  K  R   S  E  T   E  T  V  A   D  A  L

CGGCAGGGCG CACGCCCACG GCACCGGACC CGACTGGCAC GCCTGGTTCG CCGGCTCCGG GGCGCACCGC GTGGACCTGC CCACGTACTC CTTCCGGCGC 12400
 G  R  A   H  A  H  G   T  G  P   D  W  H   A  W  F  A   G  S  G   A  H  R   V  D  L  P   T  Y  S   F  R  R

GACCGCTACT GGCTGGACGC CCCGGCGGCC GACACCGCGG TGGACACCGC CGGCCTCGGT CTCGGCACCG CCGACCACCC GCTGCTCGGC GCCGTGGTCA 12500
D  R  Y  W   L  D  A   P  A  A   D  T  A  V   D  T  A   G  L  G   L  G  T  A   D  H  P   L  L  G   A  V  V  S
```

FIG. 31Y

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90       100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GCCTTCCGGA CCGGGACGGC CTGCTGCTCA CCGGCCGCCT CTCCCTGCGC ACCCACCCGT GGCTCGCGGA CCACGCCGTC CTGGGGAGCG TCCTGCTCCC 12600
  L  P  D   R  D  G   L  L  L  T   G  R  L   S  L  R   T  H  P  W   L  A  D   H  A  V   L  G  S  V   L  L  P

CGGCGCCGCG ATGGTCGAAC TCGCCGCGCA CGCTGCGGAG TCCGCCGGTC TGCGTGACGT GCGGGAGCTG ACCCTCCTTG AACCGCTGGT ACTGCCCGAG 12700
  G  A  A   M  V  E  L   A  A  H   A  A  E   S  A  G  L   R  D  V   R  E  L   T  L  L  E   P  L  V   L  P  E

CACGGTGGCG TCGAGCTGCG CGTGACGGTC GGGGCGCCGG CCGGAGAGCC CGGTGGCGAG TCGGCCGGGG ACGGCGCACG GCCCGTCTCC CTCCACTCGC 12800
  H  G  G  V   E  L  R   V  T  V   G  A  P  A   G  E  P   G  G  E   S  A  G  D   G  A  R   P  V  S   L  H  S  R

                          KpnI                                                   MscI
                       Acc65I                                                    BalI
                          ▼  ▼                                                   ▼
GGCTCGCCGA CGCGCCCGCC GGTACCGCCT GGTCCTGCCA CGCGACCGGT CTGCTGGCCA CCGACCGGCC CGAGCTTCCC GTCGCGCCCG ACCGTGCGGC 12900
  L  A  D   A  P  A   G  T  A  W   S  C  H   A  T  G   L  L  A  T   D  R  P   E  L  P   V  A  P  D   R  A  A

CATGTGGCCG CCGCAGGGCG CCGAGGAGGT GCCGCTCGAC GGTCTCTACG AGCGGCTCGA CGGGAACGGC CTCGCCTTCG GTCCGCTGTT CCAGGGGCTG 13000
  M  W  P   P  Q  G  A   E  E  V   P  L  D   G  L  Y  E   R  L  D   G  N  G   L  A  F  G   P  L  F   Q  G  L
```

FIG. 31Z

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90       100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                                          BsmI
                                                                           ▼
AACGCGGTGT GGCGGTACGA GGGTGAGGTC TTCGCCGACA TCGCGCTCCC CGCCACCACG AATGCGACCG CGCCCGCGAC CGCGAACGGC GGCGGGAGTG 13100
N  A  V  W   R  Y  E   G  E  V   F  A  D  I   A  L  P   A  T  T   N  A  T  A   P  A  T   A  N  G   G  G  S  A

CGGCGGCGGC CCCCTACGGC ATCCACCCCG CCCTGCTCGA CGCTTCGCTG CACGCCATCG CGGTCGGCGG TCTCGTCGAC GAGCCCGAGC TCGTCCGCGT 13200
  A  A  A   P  Y  G   I  H  P   A  L  L   D  A  S   L  H  A   I  A  V   G  G   L  V  D   E  P  E   L  V  R  V

                                ApaLI
                                  ▼
CCCCTTCCAC TGGAGCGGTG TCACCGTGCA CGCGGCCGGT GCCGCGGCGG CCCGGGTCCG TCTCGCCTCC GCGGGGACGG ACGCCGTCTC GCTGTCCCTG 13300
 P  F  H   W  S  G  V   T  V  H   A  A  G   A  A  A  A   R  V  R   L  A  S   A  G  T  D   A  V  S   L  S  L

ACGGACGGCG AGGGACGCCC GCTGGTCTCC GTGGAACGGC TCACGCTGCG CCCGGTCACC GCCGATCAGG CGGCGGCGAG CCGCGTCGGC GGGCTGATGC 13400
T  D  G  E   G  R  P   L  V  S   V  E  R  L   T  L  R   P  V  T   A  D  Q  A   A  A  S   R  V  G   G  L  M  H

ACCGGGTGGC CTGGCGTCCG TACGCCCTCG CCTCGTCCGG CGAACAGGAC CCGCACGCCA CTTCGTACGG GCCGACCGCC GTCCTCGGCA AGGACGAGCT 13500
  R  V  A   W  R  P   Y  A  L  A   S  S  G   E  Q  D   P  H  A   T  S  Y   G  P   T  A  V   L  G  K   D  E  L
```

FIG. 31AA

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GAAGGTCGCC GCCGCCCTGG AGTCCGCGGG CGTCGAAGTC GGGCTCTACC CCGACCTGGC CGCGCTGTCC CAGGACGTGG CGGCCGGCGC CCCGGCGCCC 13600
 K  V  A  A  A  L  E  S  A  G  V  E  V  G  L  Y  P  D  L  A  A  L  S  Q  D  V  A  A  G  A  P  A  P

CGTACCGTCC TTGCGCCGCT GCCCGCGGGT CCCGCCGACG GCGGCGCGGA GGGTGTACGG GGCACGGTGG CCCGGACGCT GGAGCTGCTC CAGGCCTGGC 13700
 R  T  V  L  A  P  L  P  A  G  P  A  D  G  G  A  E  G  V  R  G  T  V  A  R  T  L  E  L  L  Q  A  W  L

TGGCCGACGA GCACCTCGCG GGCACCCGCC TGCTCCTGGT CACCCGCGGT GCGGTGCGGG ACCCCGAGGG GTCCGGCGCC GACGATGGCG GCGAGGACCT 13800
  A  D  E  H  L  A  G  T  R  L  L  L  V  T  R  G  A  V  R  D  P  E  G  S  G  A  D  D  G  G  E  D  L

        NotI
        ▼
GTCGCACGCG GCCGCCTGGG GTCTCGTACG GACCGCGCAG ACCGAGAACC CCGGCCGCTT CGGCCTTCTC GACCTGGCCG ACGACGCCTC GTCGTACCGG 13900
 S  H  A  A  A  W  G  L  V  R  T  A  Q  T  E  N  P  G  R  F  G  L  L  D  L  A  D  D  A  S  S  Y  R

                                                                                BstXI
                                                                                ▼
ACCCTGCCGT CGGTGCTCTC CGACGCGGGC CTGCGCGACG AACCGCAGCT CGCCCTGCAC GACGGCACCA TCAGGCTGGC CCGCCTGGCC TCCGTCCGGC 14000
 T  L  P  S  V  L  S  D  A  G  L  R  D  E  P  Q  L  A  L  H  D  G  T  I  R  L  A  R  L  A  S  V  R  P
```

FIG. 31BB

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                                                          PmlI
                                                                                          ▼
CCGAGACCGG CACCGCCGCA CCGGCGCTCG CCCCGGAGGG CACGGTCCTG CTGACCGGCG GCACCGGCGG CCTGGGCGGA CTGGTCGCCC GGCACGTGGT 14100
  E  T  G  T  A  A  P  A  L  A  P  E  G  T  V  L  L  T  G  G  T  G  G  L  G  G  L  V  A  R  H  V  V

                                                                    ApaLI
                                                                    ▼
GGGCGAGTGG GGCGTACGAC GCCTGCTGCT GGTGAGCCGG CGGGGCACGG ACGCCCCGGG CGCCGACGAG CTCGTGCACG AGCTGGAGGC CCTGGGAGCC 14200
 G  E  W  G  V  R  R  L  L  L  V  S  R  R  G  T  D  A  P  G  A  D  E  L  V  H  E  L  E  A  L  G  A

GACGTCTCGG TGGCCGCGTG CGACGTCGCC GACCGCGAAG CCCTCACCGC CGTACTCGAC GCCATCCCCG CCGAACACCC GCTCACCGCG GTCGTCCACA 14300
D  V  S  V  A  A  C  D  V  A  D  R  E  A  L  T  A  V  L  D  A  I  P  A  E  H  P  L  T  A  V  V  H  T

CGGCAGGCGT CCTCTCCGAC GGCACCCTCC CGTCCATGAC GACGGAGGAC GTGGAACACG TACTGCGGCC CAAGGTCGAC GCCGCGTTCC TCCTCGACGA 14400
 A  G  V  L  S  D  G  T  L  P  S  M  T  T  E  D  V  E  H  V  L  R  P  K  V  D  A  A  F  L  L  D  E

ACTCACCTCG ACGCCCGCAT ACGACCTGGC AGCGTTCGTC ATGTTCTCCT CCGCCGCCGC CGTCTTCGGT GGCGCGGGGC AGGGCGCCTA CGCCGCCGCC 14500
 L  T  S  T  P  A  Y  D  L  A  A  F  V  M  F  S  S  A  A  A  V  F  G  G  A  G  Q  G  A  Y  A  A  A
```

FIG. 31CC

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

AACGCCACCC TCGACGCCCT CGCCTGGCGC CGCCGGGCAG CCGGACTCCC CGCCCTCTCC CTCGGCTGGG GCCTCTGGGC CGAGACCAGC GGCATGACCG 14600
N  A  T  L  D  A  L  A  W  R  R  R  A  A  G  L  P  A  L  S  L  G  W  G  L  W  A  E  T  S  G  M  T  G

GCGAGCTCGG CCAGGCGGAC CTGCGCCGGA TGAGCCGCGC GGGCATCGGC GGGATCAGCG ACGCCGAGGG CATCGCGCTC CTCGACGCCG CCCTCCGCGA 14700
  E  L  G  Q  A  D  L  R  R  M  S  R  A  G  I  G  G  I  S  D  A  E  G  I  A  L  L  D  A  A  L  R  D

CGACCGCCAC CCGGTCCTGC TGCCCCTGCG GCTCGACGCC GCCGGGCTGC GGGACGCGGC CGGGAACGAC CCGGCCGGAA TCCCGGCGCT CTTCCGGGAC 14800
 D  R  H  P  V  L  L  P  L  R  L  D  A  A  G  L  R  D  A  A  G  N  D  P  A  G  I  P  A  L  F  R  D

GTCGTCGGCG CCAGGACCGT CCGGGCCCGG CCGTCCGCGG CCTCCGCCTC GACGACAGCC GGGACGGCCG GCACGCCGGG GACGGCGGAC GGCGCGGCGG 14900
V  V  G  A  R  T  V  R  A  R  P  S  A  A  S  A  S  T  T  A  G  T  A  G  T  P  G  T  A  D  G  A  A  E

                                                                                XhoI
                                                                                PaeR7I
                                                                                ▼
AAACGGCGGC GGTCACGCTC GCCGACCGGG CCGCCACCGT GGACGGGCCC GCACGGCAGC GCCTGCTGCT CGAGTTCGTC GTCGGCGAGG TCGCCGAAGT 15000
  T  A  A  V  T  L  A  D  R  A  A  T  V  D  G  P  A  R  Q  R  L  L  L  E  F  V  V  G  E  V  A  E  V
```

FIG. 31DD

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

ScaI
▼
ACTCGGCCAC GCCCGCGGTC ACCGGATCGA CGCCGAACGG GGCTTCCTCG ACCTCGGCTT CGACTCCCTG ACCGCCGTCG AACTCCGCAA CCGGCTCAAC 15100
 L  G  H   A  R  G  H   R  I  D   A  E  R   G  F  L  D   L  G  F   D  S  L   T  A  V  E   L  R  N   R  L  N

TCCGCCGGTG GCCTCGCCCT CCCGGCGACC CTGGTCTTCG ACCACCCAAG CCCGGCGGCA CTCGCCTCCC ACCTGGACGC CGAGCTGCCG CGCGGCGCCT 15200
 S  A  G  G   L  A  L   P  A  T   L  V  F  D   H  P  S   P  A  A   L  A  S  H   L  D  A   E  L  P   R  G  A  S

CGGACCAGGA CGGAGCCGGG AACCGGAACG GGAACGAGAA CGGGACGACG GCGTCCCGGA GCACCGCCGA GACGGACGCG CTGCTGGCAC AACTGACCCG 15300
  D  Q  D   G  A  G   N  R  N  G   N  E  N   G  T  T   A  S  R  S   T  A  E   T  D  A   L  L  A  Q   L  T  R

CCTGGAAGGC GCCTTGGTGC TGACGGGCCT CTCGGACGCC CCCGGGAGCG AAGAAGTCCT GGAGCACCTG CGGTCCCTGC GCTCGATGGT CACGGGCGAG 15400
 L  E  G   A  L  V  L   T  G  L   S  D  A   P  G  S  E   E  V  L   E  H  L   R  S  L  R   S  M  V   T  G  E

ACCGGGACCG GGACCGCGTC CGGAGCCCCG GACGGCGCCG GGTCCGGCGC CGAGGACCGG CCCTGGGCGG CCGGGGACGG AGCCGGGGGC GGGAGTGAGG 15500
 T  G  T  G   T  A  S   G  A  P   D  G  A  G   S  G  A   E  D  R   P  W  A  A   G  D  G   A  G  G   G  S  E  D
```

FIG. 31EE

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                       BspHI
                         ▼
ACGGCGCGGG AGTGCCGGAC TTCATGAACG CCTCGGCCGA GGAACTCTTC GGCCTCCTCG ACCAGGACCC CAGCACGGAC TGATCCCTGC CGCACGGTCG 15600
  G  A  G   V  P  D   F  M  N  A   S  A  E   E  L  F   G  L  L  D   Q  D  P   S  T  D

                                          XmnI
                                           ▼
CCTCCCGCCC CGGACCCCGT CCCGGGCACC TCGACTCGAA TCACTTCATG CGCGCCTCGG GCGCCTCCAG GAACTCAAGG GGACAGCGTG TCCACGGTGA 15700
                                                                                 V  S  T  V  N

ACGAAGAGAA GTACCTCGAC TACCTGCGTC GTGCCACGGC GGACCTCCAC GAGGCCCGTG GCCGCCTCCG CGAGCTGGAG GCGAAGGCGG GCGAGCCGGT 15800
  E  E  K   Y  L  D   Y  L  R  R   A  T  A   D  L  H   E  A  R  G   R  L  R   E  L  E   A  K  A  G   E  P  V

GGCGATCGTC GGCATGGCCT GCCGCCTGCC CGGCGGCGTC GCCTCGCCCG AGGACCTGTG GCGGCTGGTG GCCGGCGGCG AGGACGCGAT CTCGGAGTTC 15900
 A  I  V   G  M  A  C   R  L  P   G  G  V   A  S  P  E   D  L  W   R  L  V   A  G  G  E   D  A  I   S  E  F

CCCCAGGACC GCGGCTGGGA CGTGGAGGGC CTGTACGACC CGAACCCGGA GGCCACGGGC AAGAGTTACG CCCGCGAGGC CGGATTCCTG TACGAGGCGG 16000
P  Q  D  R   G  W  D   V  E  G   L  Y  D  P   N  P  E   A  T  G   K  S  Y  A   R  E  A   G  F  L   Y  E  A  G
```

FIG. 31FF

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GCGAGTTCGA CGCCGACTTC TTCGGGATCT CGCCGCGCGA GGCCCTCGCC ATGGACCCGC AGCAGCGTCT CCTCCTGGAG GCCTCCTGGG AGGCGTTCGA 16100
  E  F  D  A  D  F  F  G  I  S  P  R  E  A  L  A  M  D  P  Q  Q  R  L  L  L  E  A  S  W  E  A  F  E

       BamHI
         ▼
GCACGCCGGG ATCCCGGCGG CCACCGCGCG CGGCACCTCG GTCGGCGTCT TCACCGGCGT GATGTACCAC GACTACGCCA CCCGTCTCAC CGATGTCCCG 16200
 H  A  G  I  P  A  A  T  A  R  G  T  S  V  G  V  F  T  G  V  M  Y  H  D  Y  A  T  R  L  T  D  V  P

GAGGGCATCG AGGGCTACCT GGGCACCGGC AACTCCGGCA GTGTCGCCTC GGGCCGCGTC GCGTACACGC TTGGCCTGGA GGGGCCGGCC GTCACGGTCG 16300
E  G  I  E  G  Y  L  G  T  G  N  S  G  S  V  A  S  G  R  V  A  Y  T  L  G  L  E  G  P  A  V  T  V  D

ACACCGCCTG CTCGTCCTCG CTGGTCGCCC TGCACCTCGC CGTGCAGGCC CTGCGCAAGG GCGAGGTCGA CATGGCGCTC GCCGGCGGCG TGACGGTCAT 16400
  T  A  C  S  S  S  L  V  A  L  H  L  A  V  Q  A  L  R  K  G  E  V  D  M  A  L  A  G  G  V  T  V  M

                                                                                  XmnI
                                                                                    ▼
GTCGACGCCC AGCACCTTCG TCGAGTTCAG CCGTCAGCGC GGGCTGGCGC CGGACGGCCG GTCGAAGTCC TTCTCGTCGA CGGCCGACGG CACCAGCTGG 16500
 S  T  P  S  T  F  V  E  F  S  R  Q  R  G  L  A  P  D  G  R  S  K  S  F  S  S  T  A  D  G  T  S  W
```

FIG. 31GG

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                                    BamHI
                                                                      ▼
TCCGAGGGCG TCGGCGTCCT CCTCGTCGAG CGCCTGTCCG ACGCGCGTCG CAAGGGCCAT CGGATCCTCG CCGTGGTCCG GGGCACCGCC GTCAACCAGG 16600
S  E  G  V   G  V  L   L  V  E   R  L  S  D   A  R  R   K  G  H   R  I  L  A   V  V  R   G  T  A   V  N  Q  D

ACGGCGCCAG CAGCGGCCTC ACGGCTCCGA ACGGGCCGTC GCAGCAGCGC GTCATCCGAC GTGCCCTGGC GGACGCCCGG CTCACGACCT CCGACGTGGA 16700
  G  A  S   S  G  L   T  A  P   N  G  P  S   Q  Q  R   V  I  R   R  A  L  A   D  A  R   L  T  T   S  D  V  D

CGTCGTCGAG GCCCACGGCA CGGGTACGCG ACTCGGCGAC CCGATCGAGG CGCAGGCCGT CATCGCCACG TACGGGCAGG GCCGTGACGG CGAACAGCCG 16800
 V  V  E   A  H  G   T  G  T   R  L  G   D  P  I   E  A  Q   A  V  I   A  T  Y   G  Q  G   R  D  G   E  Q  P

                                                                    BclI
                                                                      ▼
CTGCGCCTCG GGTCGTTGAA GTCCAACATC GGACACACCC AGGCCGCCGC CGGTGTCTCC GGCGTGATCA AGATGGTCCA GGCGATGCGC CACGGCGTCC 16900
L  R  L  G   S  L  K   S  N  I   G  H  T  Q   A  A  A   G  V  S   G  V  I  K   M  V  Q   A  M  R   H  G  V  L

          PmlI
            ▼
TGCCGAAGAC GCTCCACGTG GAGAAGCCGA CGGACCAGGT GGACTGGTCC GCGGGCGCGG TCGAGCTGCT CACCGAGGCC ATGGACTGGC CGGACAAGGG 17000
  P  K  T   L  H  V   E  K  P   T  D  Q  V   D  W  S   A  G  A  V   E  L  L   T  E  A   M  D  W  P   D  K  G
```

FIG. 31HH

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CGACGGCGGA CTGCGCAGGG CCGCGGTCTC CTCCTTCGGC GTCAGCGGGA CGAACGCGCA CGTCGTGCTC GAAGAGGCCC CGGCGGCCGA GGAGACCCCT 17100
 D  G  G   L  R  R  A   A  V  S   S  F  G   V  S  G  T   N  A  H   V  V  L   E  E  A  P   A  A  E   E  T  P

GCCTCCGAGG CGACCCCGGC CGTCGAGCCG TCGGTCGGCG CCGGCCTGGT GCCGTGGCTG GTGTCGGCGA AGACTCCGGC CGCGCTGGAC GCCCAGATCG 17200
A  S  E  A   T  P  A   V  E  P   S  V  G  A   G  L  V   P  W  L   V  S  A  K   T  P  A   A  L  D   A  Q  I  G

                                                                                    AscI
                                                                                     ▼
GACGCCTCGC CGCGTTCGCC TCGCAGGGCC GTACGGACGC CGCCGATCCG GGCGCGGTCG CTCGCGTACT GGCCGGCGGG CGCGCCGAGT TCGAGCACCG 17300
  R  L  A   A  F  A   S  Q  G  R   T  D  A   A  D  P   G  A  V  A   R  V  L   A  G  G   R  A  E  F   E  H  R

GGCCGTCGTG CTCGGCACCG GACAGGACGA TTTCGCGCAG GCGCTGACCG CTCCGGAAGG ACTGATACGC GGCACGCCCT CGGACGTGGG CCGGGTGGCG 17400
 A  V  V   L  G  T  G   Q  D  D   F  A  Q   A  L  T  A   P  E  G   L  I  R   G  T  P  S   D  V  G   R  V  A

                                                                                    XmnI
                                                                                     ▼
TTCGTGTTCC CCGGTCAGGG CACGCAGTGG GCCGGGATGG GCGCCGAACT CCTCGACGTG TCGAAGGAGT TCGCGGCGGC CATGGCCGAG TGCGAGAGCG 17500
F  V  F  P   G  Q  G   T  Q  W   A  G  M  G   A  E  L   L  D  V   S  K  E  F   A  A  A   M  A  E   C  E  S  A
```

FIG. 31II

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90       100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                          AscI
                                                           ▼
CGCTCTCCCG CTATGTCGAC TGGTCGCTGG AGGCCGTCGT CCGGCAGGCG CCGGGCGCGC CCACGCTGGA GCGGGTCGAC GTCGTCCAGC CCGTGACCTT 17600
  L  S  R  Y  V  D  W  S  L  E  A  V  V  R  Q  A  P  G  A  P  T  L  E  R  V  D  V  V  Q  P  V  T  F

CGCTGTCATG GTTTCGCTGG CGAAGGTCTG GCAGCACCAC GGCGTGACGC CGCAGGCCGT CGTCGGCCAC TCGCAGGGCG AGATCGCCGC CGCGTACGTC 17700
  A  V  M  V  S  L  A  K  V  W  Q  H  H  G  V  T  P  Q  A  V  V  G  H  S  Q  G  E  I  A  A  A  Y  V

GCCGGTGCCC TCACCCTCGA CGACGCCGCC CGCGTCGTCA CCCTGCGCAG CAAGTCCATC GCCGCCCACC TCGCCGGCAA GGGCGGCATG ATCTCCCTCG 17800
  A  G  A  L  T  L  D  D  A  A  R  V  V  T  L  R  S  K  S  I  A  A  H  L  A  G  K  G  G  M  I  S  L  A

CCCTCAGCGA GGAAGCCACC CGGCAGCGCA TCGAGAACCT CCACGGACTG TCGATCGCCG CCGTCAACGG CCCCACCGCC ACCGTGGTTT CGGGCGACCC 17900
  L  S  E  E  A  T  R  Q  R  I  E  N  L  H  G  L  S  I  A  A  V  N  G  P  T  A  T  V  V  S  G  D  P

CACCCAGATC CAAGAGCTCG CTCAGGCGTG TGAGGCCGAC GGGGTCCGCG CACGGATCAT CCCCGTCGAC TACGCCTCCC ACAGCGCCCA CGTCGAGACC 18000
  T  Q  I  Q  E  L  A  Q  A  C  E  A  D  G  V  R  A  R  I  I  P  V  D  Y  A  S  H  S  A  H  V  E  T
```

FIG. 31JJ

pikPKS Sequence

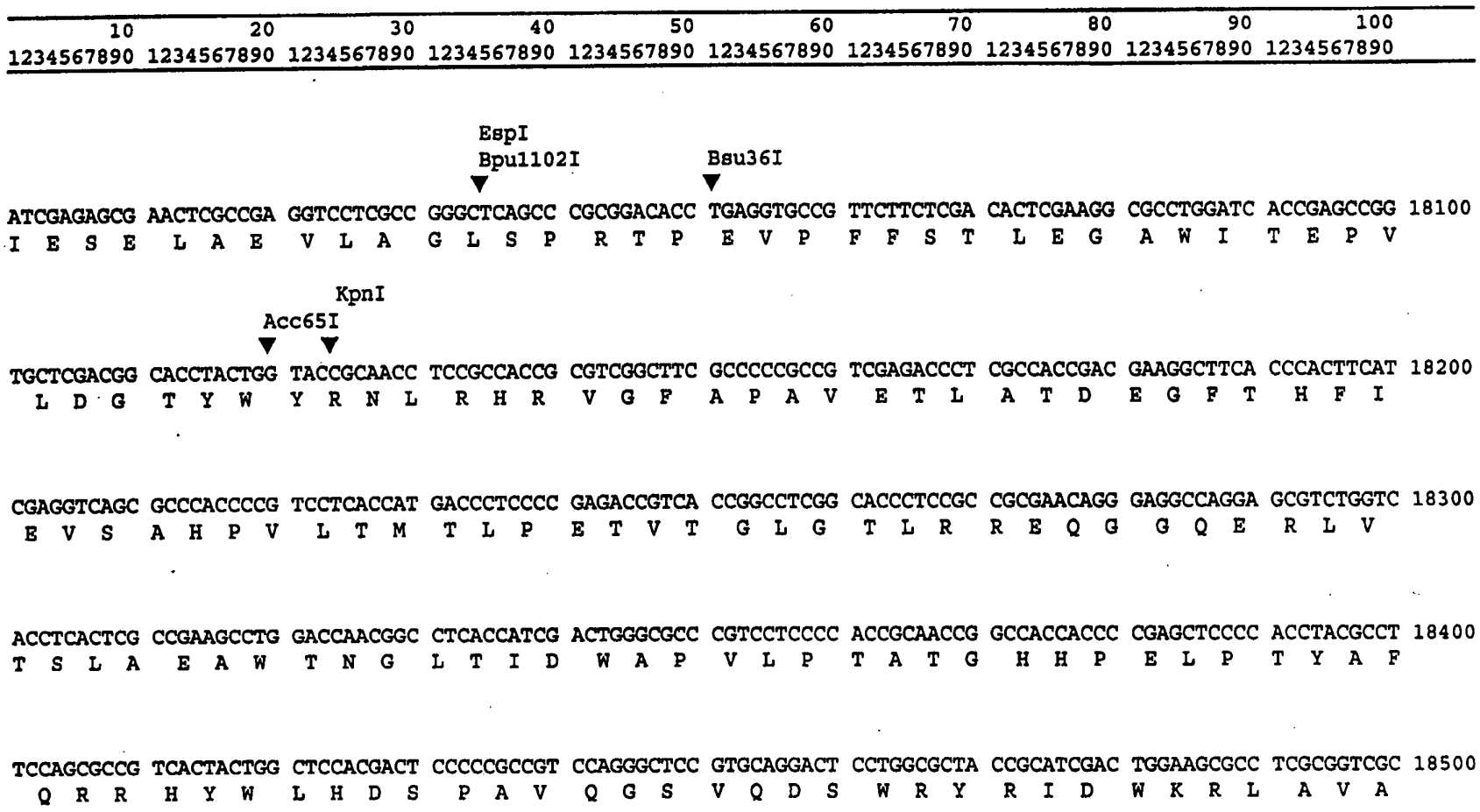

FIG. 31KK

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

   MluI
   ▼
CGACGCGTCC GAGCGCGCCG GGCTGTCCGG GCGCTGGCTC GTCGTCGTCC CCGAGGACCG TTCCGCCGAG GCCGCCCCGG TGCTCGCCGC GCTGTCCGGC 18600
 D  A  S   E  R  A  G   L  S  G   R  W  L   V  V  V  P   E  D  R   S  A  E   A  A  P  V   L  A  A   L  S  G

GCCGGCGCCG ACCCCGTACA GCTGGACGTG TCCCCGCTGG GCGACCGGCA GCGGCTCGCC GCGACGCTGG GCGAGGCCCT GGCGGCGGCC GGTGGAGCCG 18700
A  G  A  D   P  V  Q   L  D  V   S  P  L  G   D  R  Q   R  L  A   A  T  L  G   E  A  L   A  A  A   G  G  A  V

TCGACGGCGT CCTCTCGCTG CTCGCGTGGG ACGAGAGCGC GCACCCCGGC CACCCCGCCC CCTTCACCCG GGGCACCGGC GCCACCCTCA CCCTGGTGCA 18800
  D  G  V   L  S  L   L  A  W  D   E  S  A   H  P  G   H  P  A  P   F  T  R   G  T  G   A  T  L  T   L  V  Q

GGCGCTGGAG GACGCCGGCG TCGCCGCCCC GCTGTGGTGC GTGACCCACG GCGCGGTGTC CGTCGGCCGG GCCGACCACG TCACCTCCCC CGCCCAGGCC 18900
 A  L  E   D  A  G  V   A  A  P   L  W  C   V  T  H  G   A  V  S   V  G  R   A  D  H  V   T  S  P   A  Q  A

ATGGTGTGGG GCATGGGCCG GGTCGCCGCC CTGGAGCACC CCGAGCGGTG GGGCGGCCTG ATCGACCTGC CCTCGGACGC CGACCGGGCG GCCCTGGACC 19000
M  V  W  G   M  G  R   V  A  A   L  E  H  P   E  R  W   G  G  L   I  D  L  P   S  D  A   D  R  A   A  L  D  R
```

FIG. 31LL

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90       100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GCATGACCAC GGTCCTCGCC GGCGGTACGG GTGAGGACCA GGTCGCGGTA CGCGCCTCCG GGCTGCTCGC CCGCCGCCTC GTCCGCGCCT CCCTCCCGGC 19100
  M  T  T  V  L  A  G  G  T  G  E  D  Q  V  A  V  R  A  S  G  L  L  A  R  R  L  V  R  A  S  L  P  A

                                                                                  NotI
                                                                                   ▼
GCACGGCACG GCTTCGCCGT GGTGGCAGGC CGACGGCACG GTGCTCGTCA CCGGTGCCGA GGAGCCTGCG GCCGCCGAGG CCGCACGCCG GCTGGCCCGC 19200
 H  G  T  A  S  P  W  W  Q  A  D  G  T  V  L  V  T  G  A  E  E  P  A  A  A  E  A  A  R  R  L  A  R

GACGGCGCCG GACACCTCCT CCTCCACACC ACCCCCTCCG GCAGCGAAGG CGCCGAAGGC ACCTCCGGTG CCGCCGAGGA CTCCGGCCTC GCCGGGCTCG 19300
D  G  A  G  H  L  L  L  H  T  T  P  S  G  S  E  G  A  E  G  T  S  G  A  A  E  D  S  G  L  A  G  L  V

                                                                                  NotI
                                                                                   ▼
TCGCCGAACT CGCGGACCTG GGCGCGACGG CCACCGTCGT GACCTGCGAC CTCACGGACG CGGAGGCGGC CGCCCGGCTG CTCGCCGGCG TCTCCGACGC 19400
  A  E  L  A  D  L  G  A  T  A  T  V  V  T  C  D  L  T  D  A  E  A  A  A  R  L  L  A  G  V  S  D  A

       EspI
       Bpu1102I
        ▼
GCACCCGCTC AGCGCCGTCC TCCACCTGCC GCCCACCGTC GACTCCGAGC CGCTCGCCGC GACCGACGCG GACGCGCTCG CCCGTGTCGT GACCGCGAAG 19500
 H  P  L  S  A  V  L  H  L  P  P  T  V  D  S  E  P  L  A  A  T  D  A  D  A  L  A  R  V  V  T  A  K
```

FIG. 31MM

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GCCACCGCCG CGCTCCACCT GGACCGCCTC CTGCGGGAGG CCGCGGCTGC CGGAGGCCGT CCGCCCGTCC TGGTCCTCTT CTCCTCGGTC GCCGCGATCT 19600
A  T  A  A  L  H  L  D  R  L  L  R  E  A  A  A  A  G  G  R  P  P  V  L  V  L  F  S  S  V  A  A  I  W

GGGGCGGCGC CGGTCAGGGC GCGTACGCCG CCGGTACGGC CTTCCTCGAC GCCCTCGCCG GTCAGCACCG GGCCGACGGC CCCACCGTGA CCTCGGTGGC 19700
  G  G  A  G  Q  G  A  Y  A  A  G  T  A  F  L  D  A  L  A  G  Q  H  R  A  D  G  P  T  V  T  S  V  A

CTGGAGCCCC TGGGAGGGCA GCCGCGTCAC CGAGGGTGCG ACCGGGGAGC GGCTGCGCCG CCTCGGCCTG CGCCCCCTCG CCCCCGCGAC GGCGCTCACC 19800
 W  S  P  W  E  G  S  R  V  T  E  G  A  T  G  E  R  L  R  R  L  G  L  R  P  L  A  P  A  T  A  L  T

GCCCTGGACA CCGCGCTCGG CCACGGCGAC ACCGCCGTCA CGATCGCCGA CGTCGACTGG TCGAGCTTCG CCCCCGGCTT CACCACGGCC CGGCCGGGCA 19900
A  L  D  T  A  L  G  H  G  D  T  A  V  T  I  A  D  V  D  W  S  S  F  A  P  G  F  T  T  A  R  P  G  T

                                AscI
                                 ▼
CCCTCCTCGC CGATCTGCCC GAGGCGCGCC GCGCGCTCGA CGAGCAGCAG TCGACGACGG CCGCCGACGA CACCGTCCTG AGCCGCGAGC TCGGTGCGCT 20000
  L  L  A  D  L  P  E  A  R  R  A  L  D  E  Q  Q  S  T  T  A  A  D  D  T  V  L  S  R  E  L  G  A  L
```

FIG. 31NN

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CACCGGCGCC GAACAGCAGC GCCGTATGCA GGAGTTGGTC CGCGAGCACC TCGCCGTGGT CCTCAACCAC CCCTCCCCCG AGGCCGTCGA CACGGGGCGG 20100
 T  G  A  E  Q  Q  R  R  M  Q  E  L  V  R  E  H  L  A  V  V  L  N  H  P  S  P  E  A  V  D  T  G  R

                                                                                  BstXI
                                                                                    ▼
GCCTTCCGTG ACCTCGGATT CGACTCGCTG ACGGCGGTCG AGCTCCGCAA CCGCCTCAAG AACGCCACCG GCCTGGCCCT CCCGGCCACT CTGGTCTTCG 20200
 A  F  R  D  L  G  F  D  S  L  T  A  V  E  L  R  N  R  L  K  N  A  T  G  L  A  L  P  A  T  L  V  F  D

ACTACCCGAC CCCCCGGACG CTGGCGGAGT TCCTCCTCGC GGAGATCCTG GGCGAGCAGG CCGGTGCCGG CGAGCAGCTT CCGGTGGACG GCGGGGTCGA 20300
  Y  P  T  P  R  T  L  A  E  F  L  L  A  E  I  L  G  E  Q  A  G  A  G  E  Q  L  P  V  D  G  G  V  D

CGACGAGCCC GTCGCGATCG TCGGCATGGC GTGCCGCCTG CCGGGCGGTG TCGCCTCGCC GGAGGACCTG TGGCGGCTGG TGGCCGGCGG CGAGGACGCG 20400
 D  E  P  V  A  I  V  G  M  A  C  R  L  P  G  G  V  A  S  P  E  D  L  W  R  L  V  A  G  G  E  D  A

                                                                                  MluI
                                                                                    ▼
ATCTCCGGCT TCCCGCAGGA CCGCGGCTGG GACGTGGAGG GGCTGTACGA CCCGGACCCG GACGCGTCCG GGCGGACGTA CTGCCGTGCC GGTGGCTTCC 20500
 I  S  G  F  P  Q  D  R  G  W  D  V  E  G  L  Y  D  P  D  P  D  A  S  G  R  T  Y  C  R  A  G  G  F  L
```

FIG. 3100

pikPKS Sequence

```
      10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

TCGACGAGGC GGGCGAGTTC GACGCCGACT TCTTCGGGAT CTCGCCGCGC GAGGCCCTCG CCATGGACCC GCAGCAGCGG CTCCTCCTGG AGACCTCCTG 20600
  D  E  A  G  E  F  D  A  D  F  F  G  I  S  P  R  E  A  L  A  M  D  P  Q  Q  R  L  L  L  E  T  S  W

GGAGGCCGTC GAGGACGCCG GGATCGACCC GACCTCCCTT CAGGGGCAGC AGGTCGGCGT GTTCGCGGGC ACCAACGGCC CCCACTACGA GCCGCTGCTC 20700
  E  A  V  E  D  A  G  I  D  P  T  S  L  Q  G  Q  Q  V  G  V  F  A  G  T  N  G  P  H  Y  E  P  L  L

CGCAACACCG CCGAGGATCT TGAGGGTTAC GTCGGGACGG GCAACGCCGC CAGCATCATG TCGGGCCGTG TCTCGTACAC CCTCGGCCTG GAGGGCCCGG 20800
  R  N  T  A  E  D  L  E  G  Y  V  G  T  G  N  A  A  S  I  M  S  G  R  V  S  Y  T  L  G  L  E  G  P  A

                                                                                                  BsmI
                                                                                                   ▼
CCGTCACGGT CGACACCGCC TGCTCCTCCT CGCTGGTCGC CCTGCACCTC GCCGTGCAGG CCCTGCGCAA GGGCGAATGC GGACTGGCGC TCGCGGGCGG 20900
  V  T  V  D  T  A  C  S  S  S  L  V  A  L  H  L  A  V  Q  A  L  R  K  G  E  C  G  L  A  L  A  G  G

                                                                                                 XmnI
                                                                                                   ▼
TGTGACGGTC ATGTCGACGC CCACGACGTT CGTGGAGTTC AGCCGGCAGC GCGGGCTCGC GGAGGACGGC CGGTCGAAGG CGTTCGCCGC GTCGGCGGAC 21000
  V  T  V  M  S  T  P  T  T  F  V  E  F  S  R  Q  R  G  L  A  E  D  G  R  S  K  A  F  A  A  S  A  D
```

FIG. 31PP

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                    SphI
                                    ▼
GGCTTCGGCC CGGCGGAGGG CGTCGGCATG CTCCTCGTCG AGCGCCTGTC GGACGCCCGC CGCAACGGAC ACCGTGTGCT GGCGGTCGTG CGCGGCAGCG 21100
G  F  G  P   A  E  G   V  G  M   L  L  V  E   R  L  S   D  A  R   R  N  G  H   R  V  L   A  V  V   R  G  S  A

CGGTCAACCA GGACGGCGCG AGCAACGGCC TGACCGCCCC GAACGGGCCC TCGCAGCAGC GCGTCATCCG GCGCGCGCTC GCGGACGCCC GACTGACGAC 21200
  V  N  Q   D  G  A   S  N  G   L  T  A   P  N  G   P  S  Q   Q  R  V   I  R  R   A  L  A   D  A  R   L  T  T

CGCCGACGTG GACGTCGTCG AGGCCCACGG CACGGGCACG CGACTCGGCG ACCCGATCGA GGCACAGGCC CTCATCGCCA CCTACGGCCA GGGGCGCGAC 21300
 A  D  V   D  V  V  E   A  H  G   T  G  T   R  L  G  D   P  I  E   A  Q  A   L  I  A  T   Y  G  Q   G  R  D

ACCGAACAGC CGCTGCGCCT GGGGTCGTTG AAGTCCAACA TCGGACACAC CCAGGCCGCC GCCGGTGTCT CCGGCATCAT CAAGATGGTC CAGGCGATGC 21400
T  E  Q  P   L  R  L   G  S  L   K  S  N  I   G  H  T   Q  A  A   A  G  V  S   G  I  I   K  M  V   Q  A  M  R

                                    PmlI
                                    ▼
GCCACGGCGT CCTGCCGAAG ACGCTCCACG TGGACCGGCC GTCGGACCAG ATCGACTGGT CGGCGGGCAC GGTCGAGCTG CTCACCGAGG CCATGGACTG 21500
  H  G  V   L  P  K   T  L  H  V   D  R  P   S  D  Q   I  D  W  S   A  G  T   V  E  L   L  T  E  A   M  D  W
```

FIG. 31QQ

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GCCGAGGAAG CAGGAGGGCG GGCTGCGCCG CGCGGCCGTC TCCTCCTTCG GCATCAGCGG CACGAACGCG CACATCGTGC TCGAAGAAGC CCCGGTCGAC 21600
 P  R  K   Q  E  G  G   L  R  R   A  A  V   S  S  F  G   I  S  G   T  N  A   H  I  V  L   E  E  A   P  V  D

GAGGACGCCC CGGCGGACGA GCCGTCGGTC GGCGGTGTGG TGCCGTGGCT CGTGTCCGCG AAGACTCCGG CCGCGCTGGA CGCCCAGATC GGACGCCTCG 21700
 E  D  A  P   A  D  E   P  S  V   G  G  V  V   P  W  L   V  S  A   K  T  P  A   A  L  D   A  Q  I   G  R  L  A

CCGCGTTCGC CTCGCAGGGC CGTACGGACG CCGCCGATCC GGGCGCGGTC GCTCGCGTAC TGGCCGGCGG GCGTGCGCAG TTCGAGCACC GGGCCGTCGC 21800
  A  F  A   S  Q  G   R  T  D  A   A  D  P   G  A  V   A  R  V  L   A  G  G   R  A  Q   F  E  H  R   A  V  A

                          NotI                    Bsu36I
                           ▼                       ▼
GCTCGGCACC GGACAGGACG ACCTGGCGGC CGCACTGGCC GCGCCTGAGG GTCTGGTCCG GGGTGTGGCC TCCGGTGTGG GTCGAGTGGC GTTCGTGTTC 21900
 L  G  T   G  Q  D  D   L  A  A   A  L  A   A  P  E  G   L  V  R   G  V  A   S  G  V  G   R  V  A   F  V  F

                                                        XmnI
                                                         ▼
CCGGGACAGG GCACGCAGTG GGCCGGGATG GGTGCCGAAC TCCTCGACGT GTCGAAGGAG TTCGCGGCGG CCATGGCCGA GTGCGAGGCC GCGCTCGCTC 22000
 P  G  Q  G   T  Q  W   A  G  M   G  A  E  L   L  D  V   S  K  E   F  A  A  A   M  A  E   C  E  A   A  L  A  P
```

FIG. 31RR

pikPKS Sequence

```
          10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                  AscI
                                                   ▼
CGTACGTGGA CTGGTCGCTG GAGGCCGTCG TCCGACAGGC CCCCGGCGCG CCCACGCTGG AGCGGGTCGA TGTCGTCCAG CCCGTGACGT TCGCCGTCAT 22100
  Y  V  D   W  S  L   E  A  V  V   R  Q  A   P  G  A   P  T  L  E   R  V  D   V  V  Q   P  V  T  F   A  V  M

GGTCTCGCTG GCGAAGGTCT GGCAGCACCA CGGGGTGACC CCGCAAGCCG TCGTCGGCCA CTCGCAGGGC GAGATCGCCG CCGCGTACGT CGCCGGTGCC 22200
  V  S  L   A  K  V  W   Q  H  H   G  V  T   P  Q  A   V  V  G  H   S  Q  G   E  I  A  A   A  Y  V   A  G  A

                                                                                 SphI          EspI
                                                                                  ▼          Bpu1102I
                                                                                                  ▼
CTGAGCCTGG ACGACGCCGC TCGTGTCGTG ACCCTGCGCA GCAAGTCCAT CGGCGCCCAC CTCGCGGGCC AGGGCGGCAT GCTGTCCCTC GCGCTGAGCG 22300
  L  S  L   D  D  A  A   R  V  V   T  L  R   S  K  S  I   G  A  H   L  A  G  Q   G  G  M   L  S  L   A  L  S  E

AGGCGGCCGT TGTGGAGCGA CTGGCCGGGT TCGACGGGCT GTCCGTCGCC GCCGTCAACG GGCCTACCGC CACCGTGGTT TCGGGCGACC CGACCCAGAT 22400
  A  A  V   V  E  R   L  A  G  F   D  G  L   S  V  A   A  V  N  G   P  T  A   T  V  V   S  G  D  P   T  Q  I

CCAAGAGCTC GCTCAGGCGT GTGAGGCCGA CGGGGTCCGC GCACGGATCA TCCCCGTCGA CTACGCCTCC CACAGCGCCC ACGTCGAGAC CATCGAGAGC 22500
  Q  E  L   A  Q  A  C   E  A  D   G  V  R   A  R  I  I   P  V  D   Y  A  S   H  S  A  H   V  E  T   I  E  S
```

FIG. 31SS

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GAACTCGCCG ACGTCCTGGC GGGGTTGTCC CCCCAGACAC CCCAGGTCCC CTTCTTCTCC ACCCTCGAAG GCGCCTGGAT CACCGAACCC GCCCTCGACG 22600
E  L  A  D   V  L  A   G  L  S   P  Q  T  P   Q  V  P   F  F  S   T  L  E  G   A  W  I   T  E  P   A  L  D  G

                  KpnI                                                                  MscI
                Acc65I                                                                  BalI
                  ▼  ▼                                                                  ▼
GCGGCTACTG GTACCGCAAC CTCCGCCATC GTGTGGGCTT CGCCCCGGCC GTCGAAACCC TGGCCACCGA CGAAGGCTTC ACCCACTTCG TCGAGGTCAG 22700
  G  Y  W   Y  R  N   L  R  H  R   V  G  F   A  P  A   V  E  T  L   A  T  D   E  G  F   T  H  F  V   E  V  S

CGCCCACCCC GTCCTCACCA TGGCGCTGCC CGAGACCGTC ACCGGACTCG GCACCCTCCG CCGTGACAAC GGCGGACAGC ACCGCCTCAC CACCTCCCTC 22800
 A  H  P   V  L  T  M   A  L  P   E  T  V   T  G  L  G   T  L  R   R  D  N   G  G  Q  H   R  L  T   T  S  L

GCCGAGGCCT GGGCCAACGG CCTCACCGTC GACTGGGCCT CTCTCCTCCC CACCACGACC ACCCACCCCG ATCTGCCCAC CTACGCCTTC CAGACCGAGC 22900
A  E  A  W   A  N  G   L  T  V   D  W  A  S   L  L  P   T  T  T   T  H  P  D   L  P  T   Y  A  F   Q  T  E  R

Eco47III
▼
GCTACTGGCC GCAGCCCGAC CTCTCCGCCG CCGGTGACAT CACCTCCGCC GGTCTCGGGG CGGCCGAGCA CCCGCTGCTC GGCGCGGCCG TGGCGCTCGC 23000
  Y  W  P   Q  P  D   L  S  A  A   G  D  I   T  S  A   G  L  G  A   A  E  H   P  L  L   G  A  A  V   A  L  A
```

FIG. 31TT

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GGACTCCGAC GGCTGCCTGC TCACGGGGAG CCTCTCCCTC CGTACGCACC CCTGGCTGGC GGACCACGCG GTGGCCGGCA CCGTGCTGCT GCCGGGAACG 23100
 D  S  D   G  C  L  L   T  G  S   L  S  L   R  T  H  P   W  L  A   D  H  A   V  A  G  T   V  L  L   P  G  T

GCGTTCGTGG AGCTGGCGTT CCGAGCCGGG GACCAGGTCG GTTGCGATCT GGTCGAGGAG CTCACCCTCG ACGCGCCGCT CGTGCTGCCC CGTCGTGGCG 23200
A  F  V  E   L  A  F   R  A  G   D  Q  V  G   C  D  L   V  E  E   L  T  L  D   A  P  L   V  L  P   R  R  G  A

CGGTCCGTGT GCAGCTGTCC GTCGGCGCGA GCGACGAGTC CGGGCGTCGT ACCTTCGGGC TCTACGCGCA CCCGGAGGAC GCGCCGGGCG AGGCGGAGTG 23300
 V  R  V   Q  L  S   V  G  A  S   D  E  S   G  R  R   T  F  G  L   Y  A  H   P  E  D   A  P  G  E   A  E  W

GACGCGGCAC GCCACCGGTG TGCTGGCCGC CCGTGCGGAC CGCACCGCCC CCGTCGCCGA CCCGGAGGCC TGGCCGCCGC CGGGCGCCGA GCCGGTGGAC 23400
 T  R  H   A  T  G  V   L  A  A   R  A  D   R  T  A  P   V  A  D   P  E  A   W  P  P  P   G  A  E   P  V  D

                    Eco47III
                       ▼
GTGGACGGTC TGTACGAGCG CTTCGCGGCG AACGGCTACG GCTACGGCCC CCTCTTCCAG GGCGTCCGTG GTGTCTGGCG GCGTGGCGAC GAGGTGTTCG 23500
V  D  G  L   Y  E  R   F  A  A   N  G  Y  G   Y  G  P   L  F  Q   G  V  R  G   V  W  R   R  G  D   E  V  F  A
```

FIG. 31UU

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CCGACGTGGC CCTGCCGGCC GAGGTCGCCG GTGCCGAGGG CGCGCGGTTC GGCCTTCACC CGGCGCTGCT CGACGCCGCC GTGCAGGCGG CCGGTGCGGG 23600
  D  V  A  L  P  A  E  V  A  G  A  E  G  A  R  F  G  L  H  P  A  L  L  D  A  A  V  Q  A  A  G  A  G

CCGGGGCGTT CGGCGCGGGC ACGCGGCTGC CGTTCGCCTG GAGCGGGATC TCCTGTACGC GGTCGGCGCC ACCGCCCTCC GCGTGCGGCT GGCCCCCGCC 23700
  R  G  V  R  R  G  H  A  A  A  V  R  L  E  R  D  L  L  Y  A  V  G  A  T  A  L  R  V  R  L  A  P  A

GGCCCGGACA CGGTGTCCGT GAGCGCCGCC GACTCCTCCG GGCAGCCGGT GTTCGCCGCG GACTCCCTCA CGGTGCTGCC CGTCGACCCC GCGCAGCTGG 23800
  G  P  D  T  V  S  V  S  A  A  D  S  S  G  Q  P  V  F  A  A  D  S  L  T  V  L  P  V  D  P  A  Q  L  A

CGGCCTTCAG CGACCCGACT CTGGACGCGC TGCACCTGCT GGAGTGGACC GCCTGGGACG GTGCCGCGCA GGCCCTGCCC GGCGCGGTCG TGCTGGGCGG 23900
  A  F  S  D  P  T  L  D  A  L  H  L  L  E  W  T  A  W  D  G  A  A  Q  A  L  P  G  A  V  V  L  G  G

CGACGCCGAC GGTCTCGCCG CGGCGCTGCG CGCCGGTGGC ACCGAGGTCC TGTCCTTCCC GGACCTTACG GACCTGGTGG AGGCCGTCGA CCGGGGCGAG 24000
  D  A  D  G  L  A  A  A  L  R  A  G  G  T  E  V  L  S  F  P  D  L  T  D  L  V  E  A  V  D  R  G  E
```

FIG. 31W

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

ACCCCGGCCC CGGCGACCGT CCTGGTGGCC TGCCCCGCCG CCGGCCCCGA TGGGCCGGAG CATGTCCGCG AGGCCCTGCA CGGGTCGCTC GCGCTGATGC 24100
T  P  A  P   A  T  V   L  V  A   C  P  A  A   G  P  D   G  P  E   H  V  R  E   A  L  H   G  S  L   A  L  M  Q

AGGCCTGGCT GGCCGACGAG CGGTTCACCG ATGGGCGCCT GGTGCTCGTG ACCCGCGACG CGGTCGCCGC CCGTTCCGGC GACGGCCTGC GGTCCACGGG 24200
  A  W  L   A  D  E   R  F  T  D   G  R  L   V  L  V   T  R  D  A   V  A  A   R  S  G   D  G  L  R   S  T  G

ACAGGCCGCC GTCTGGGGCC TCGGCCGGTC CGCGCAGACG GAGAGCCCGG GCCGGTTCGT CCTGCTCGAC CTCGCCGGGG AAGCCCGGAC GGCCGGGGAC 24300
 Q  A  A   V  W  G  L   G  R  S   A  Q  T   E  S  P  G   R  F  V   L  L  D   L  A  G  E   A  R  T   A  G  D

GCCACCGCCG GGGACGGCCT GACGACCGGG GACGCCACCG TCGGCGGCAC CTCTGGAGAC GCCGCCCTCG GCAGCGCCCT CGCGACCGCC CTCGGCTCGG 24400
A  T  A  G   D  G  L   T  T  G   D  A  T  V   G  G  T   S  G  D   A  A  L  G   S  A  L   A  T  A   L  G  S  G

GCGAGCCGCA GCTCGCCCTC CGGGACGGGG CGCTCCTCGT ACCCCGCCTG GCGCGGGCCG CCGCGCCCGC CGCGGCCGAC GGCCTCGCCG CGGCCGACGG 24500
  E  P  Q   L  A  L   R  D  G  A   L  L  V   P  R  L   A  R  A  A   A  P  A   A  A  D   G  L  A  A   A  D  G
```

FIG. 31WW

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CCTCGCCGCT CTGCCGCTGC CCGCCGCTCC GGCCCTCTGG CGTCTGGAGC CCGGTACGGA CGGCAGCCTG GAGAGCCTCA CGGCGGCGCC CGGCGACGCC 24600
 L  A  A  L  P  L  P  A  A  P  A  L  W  R  L  E  P  G  T  D  G  S  L  E  S  L  T  A  A  P  G  D  A

GAGACCCTCG CCCCGGAGCC GCTCGGCCCG GGACAGGTCC GCATCGCGAT CCGGGCCACC GGTCTCAACT TCCGCGACGT CCTGATCGCC CTCGGCATGT 24700
 E  T  L  A  P  E  P  L  G  P  G  Q  V  R  I  A  I  R  A  T  G  L  N  F  R  D  V  L  I  A  L  G  M  Y

ACCCCGATCC GGCGCTGATG GGCACCGAGG GAGCCGGCGT GGTCACCGCG ACCGGCCCCG GCGTCACGCA CCTCGCCCCC GGCGACCGGG TCATGGGCCT 24800
 P  D  P  A  L  M  G  T  E  G  A  G  V  V  T  A  T  G  P  G  V  T  H  L  A  P  G  D  R  V  M  G  L

GCTCTCCGGC GCGTACGCCC CGGTCGTCGT GGCGGACGCG CGGACCGTCG CGCGGATGCC CGAGGGGTGG ACGTTCGCCC AGGGCGCCTC CGTGCCGGTG 24900
 L  S  G  A  Y  A  P  V  V  V  A  D  A  R  T  V  A  R  M  P  E  G  W  T  F  A  Q  G  A  S  V  P  V

GTGTTCCTGA CGGCCGTCTA CGCCCTGCGC GACCTGGCGG ACGTCAAGCC CGGCGAGCGC CTCCTGGTCC ACTCCGCCGC CGGTGGCGTG GGCATGGCCG 25000
 V  F  L  T  A  V  Y  A  L  R  D  L  A  D  V  K  P  G  E  R  L  L  V  H  S  A  A  G  G  V  G  M  A  A
```

FIG. 31XX

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CCGTGCAGCT CGCCCGGCAC TGGGGCGTGG AGGTCCACGG CACGGCGAGT CACGGGAAGT GGGACGCCCT GCGCGCGCTC GGCCTGGACG ACGCGCACAT 25100
  V  Q  L   A  R  H   W  G  V  E   V  H  G   T  A  S   H  G  K  W   D  A  L   R  A  L   G  L  D  D   A  H  I

CGCCTCCTCC CGCACCCTGG ACTTCGAGTC CGCGTTCCGT GCCGCTTCCG GCGGGGCGGG CATGGACGTC GTACTGAACT CGCTCGCCCG CGAGTTCGTC 25200
  A  S  S   R  T  L  D   F  E  S   A  F  R   A  A  S  G   G  A  G   M  D  V   V  L  N  S   L  A  R   E  F  V

GACGCCTCGC TGCGCCTGCT CGGGCCGGGC GGCCGGTTCG TGGAGATGGG GAAGACCGAC GTCCGCGACG CGGAGCGGGT CGCCGCCGAC CACCCCGGTG 25300
  D  A  S  L   R  L  L   G  P  G   G  R  F  V   E  M  G   K  T  D   V  R  D  A   E  R  V   A  A  D   H  P  G  V

TCGGCTACCG CGCCTTCGAC CTGGGCGAGG CCGGGCCGGA GCGGATCGGC GAGATGCTCG CCGAGGTCAT CGCCCTCTTC GAGGACGGGG TGCTCCGGCA 25400
  G  Y  R   A  F  D   L  G  E  A   G  P  E   R  I  G   E  M  L  A   E  V  I   A  L  F   E  D  G  V   L  R  H

CCTGCCCGTC ACGACCTGGG ACGTGCGCCG GGCCCGCGAC GCCTTCCGGC ACGTCAGCCA GGCCCGCCAC ACGGGCAAGG TCGTCCTCAC GATGCCGTCG 25500
  L  P  V   T  T  W  D   V  R  R   A  R  D   A  F  R  H   V  S  Q   A  R  H   T  G  K  V   V  L  T   M  P  S
```

FIG. 31YY

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                                                  Pm1I
                                                                                   ▼
GGCCTCGACC CGGAGGGTAC GGTCCTGCTG ACCGGCGGCA CCGGTGCGCT GGGGGGCATC GTGGCCCGGC ACGTGGTGGG CGAGTGGGGC GTACGACGCC 25600
G  L  D  P   E  G  T   V  L  L   T  G  G  T   G  A  L   G  G  I   V  A  R  H   V  V  G   E  W  G   V  R  R  L

                                                  ApaLI
                                                   ▼
TGCTGCTCGT GAGCCGGCGG GGCACGGACG CCCCGGGCGC CGGCGAGCTC GTGCACGAGC TGGAGGCCCT GGGAGCCGAC GTCTCGGTGG CCGCGTGCGA 25700
  L  L  V   S  R  R   G  T  D  A   P  G  A   G  E  L   V  H  E  L   E  A  L   G  A  D   V  S  V  A   A  C  D

CGTCGCCGAC CGCGAAGCCC TCACCGCCGT ACTCGACTCG ATCCCCGCCG AACACCCGCT CACCGCGGTC GTCCACACGG CAGGCGTCCT CTCCGACGGC 25800
 V  A  D   R  E  A  L   T  A  V   L  D  S   I  P  A  E   H  P  L   T  A  V   V  H  T  A   G  V  L   S  D  G

ACCCTCCCCT CGATGACAGC GGAGGATGTG GAACACGTAC TGCGTCCCAA GGTCGACGCC GCGTTCCTCC TCGACGAACT CACCTCGACG CCCGGCTACG 25900
T  L  P  S   M  T  A   E  D  V   E  H  V  L   R  P  K   V  D  A   A  F  L  L   D  E  L   T  S  T   P  G  Y  D

ACCTGGCAGC GTTCGTCATG TTCTCCTCCG CCGCCGCCGT CTTCGGTGGC GCGGGGCAGG GCGCCTACGC CGCCGCCAAC GCCACCCTCG ACGCCCTCGC 26000
  L  A  A   F  V  M   F  S  S  A   A  A  V   F  G  G   A  G  Q  G   A  Y  A   A  A  N   A  T  L  D   A  L  A
```

FIG. 31ZZ

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CTGGCGCCGC CGGACAGCCG GACTCCCCGC CCTCTCCCTC GGCTGGGGCC TCTGGGCCGA GACCAGCGGC ATGACCGGCG GACTCAGCGA CACCGACCGC 26100
 W  R  R  R  T  A  G  L  P  A  L  S  L  G  W  G  L  W  A  E  T  S  G  M  T  G  G  L  S  D  T  D  R

TCGCGGCTGG CCCGTTCCGG GGCGACGCCC ATGGACAGCG AGCTGACCCT GTCCCTCCTG GACGCGGCCA TGCGCCGCGA CGACCCGGCG CTCGTCCCGA 26200
 S  R  L  A  R  S  G  A  T  P  M  D  S  E  L  T  L  S  L  L  D  A  A  M  R  R  D  D  P  A  L  V  P  I

                                                  SphI       EspI                                    AscI
                                                             Bpu1102I
TCGCCCTGGA CGTCGCCGCG CTCCGCGCCC AGCAGCGCGA CGGCATGCTG GCGCCGCTGC TCAGCGGGCT CACCCGCGGA TCGCGGGTCG GCGGCGCGCC 26300
  A  L  D  V  A  A  L  R  A  Q  Q  R  D  G  M  L  A  P  L  L  S  G  L  T  R  G  S  R  V  G  G  A  P

GGTCAACCAG CGCAGGGCAG CCGCCGGAGG CGCGGGCGAG GCGGACACGG ACCTCGGCGG GCGGCTCGCC GCGATGACAC CGGACGACCG GGTCGCGCAC 26400
 V  N  Q  R  R  A  A  A  G  G  A  G  E  A  D  T  D  L  G  G  R  L  A  A  M  T  P  D  D  R  V  A  H

                       PmlI
CTGCGGGACC TCGTCCGTAC GCACGTGGCG ACCGTCCTGG GACACGGCAC CCCGAGCCGG GTGGACCTGG AGCGGGCCTT CCGCGACACC GGTTTCGACT 26500
 L  R  D  L  V  R  T  H  V  A  T  V  L  G  H  G  T  P  S  R  V  D  L  E  R  A  F  R  D  T  G  F  D  S
```

FIG. 31AAA

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CGCTCACCGC CGTCGAACTC CGCAACCGTC TCAACGCCGC GACCGGGCTG CGGCTGCCGG CCACGCTGGT CTTCGACCAC CCCACCCCGG GGGAGCTCGC 26600
  L  T  A  V  E  L  R  N  R  L  N  A  A  T  G  L  R  L  P  A  T  L  V  F  D  H  P  T  P  G  E  L  A

CGGGCACCTG CTCGACGAAC TCGCCACGGC CGCGGGCGGG TCCTGGGCGG AAGGCACCGG GTCCGGAGAC ACGGCCTCGG CGACCGATCG GCAGACCACG 26700
  G  H  L  L  D  E  L  A  T  A  A  G  G  S  W  A  E  G  T. G  S  G  D  T  A  S  A  T  D  R  Q  T  T

GCGGCCCTCG CCGAACTCGA CCGGCTGGAA GGCGTGCTCG CCTCCCTCGC GCCCGCCGCC GGCGGCCGTC CGGAGCTCGC CGCCCGGCTC AGGGCGCTGG 26800
A  A  L  A  E  L  D  R  L  E  G  V  L  A  S  L  A  P  A  A  G  G  R  P  E  L  A  A  R  L  R  A  L  A

                                        BstXI
                                          ▼
CCGCGGCCCT GGGGGACGAC GGCGACGACG CCACCGACCT GGACGAGGCG TCCGACGACG ACCTCTTCTC CTTCATCGAC AAGGAGCTGG GCGACTCCGA 26900
  A  A  L  G  D  D  G  D  D  A  T  D  L  D  E  A  S  D  D  D  L  F  S  F  I  D  K  E  L  G  D  S  D

CTTCTGACCT GCCCGACACC ACCGGCACCA CCGGCACCAC CAGCCCCCCT CACACACGGA ACACGGAACG GACAGGCGAG AACGGGAGCC ATGGCGAACA 27000
F                                                                                                 M  A  N  N
```

FIG. 31BBB

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                  PstI
                                                  SfcI
                                                  ▼  ▼
ACGAAGACAA GCTCCGCGAC TACCTCAAGC GCGTCACCGC CGAGCTGCAG CAGAACACCA GGCGTCTGCG CGAGATCGAG GGACGCACGC ACGAGCCGGT 27100
  E  D  K  L  R  D  Y  L  K  R  V  T  A  E  L  Q  Q  N  T  R  R  L  R  E  I  E  G  R  T  H  E  P  V

GGCGATCGTG GGCATGGCCT GCCGCCTGCC GGGCGGTGTC GCCTCGCCCG AGGACCTGTG GCAGCTGGTG GCCGGGGACG GGGACGCGAT CTCGGAGTTC 27200
  A  I  V  G  M  A  C  R  L  P  G  G  V  A  S  P  E  D  L  W  Q  L  V  A  G  D  G  D  A  I  S  E  F

                                                  MluI
                                                  ▼
CCGCAGGACC GCGGCTGGGA CGTGGAGGGG CTGTACGACC CCGACCCGGA CGCGTCCGGC AGGACGTACT GCCGGTCCGG CGGATTCCTG CACGACGCCG 27300
P  Q  D  R  G  W  D  V  E  G  L  Y  D  P  D  P  D  A  S  G  R  T  Y  C  R  S  G  G  F  L  H  D  A  G

GCGAGTTCGA CGCCGACTTC TTCGGGATCT CGCCGCGCGA GGCCCTCGCC ATGGACCCGC AGCAGCGACT GTCCCTCACC ACCGCGTGGG AGGCGATCGA 27400
  E  F  D  A  D  F  F  G  I  S  P  R  E  A  L  A  M  D  P  Q  Q  R  L  S  L  T  T  A  W  E  A  I  E

GAGCGCGGGC ATCGACCCGA CGGCCCTGAA GGGCAGCGGC CTCGGCGTCT TCGTCGGCGG CTGGCACACC GGCTACACCT CGGGGCAGAC CACCGCCGTG 27500
  S  A  G  I  D  P  T  A  L  K  G  S  G  L  G  V  F  V  G  G  W  H  T  G  Y  T  S  G  Q  T  T  A  V
```

FIG. 31CCC

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CAGTCGCCCG AGCTGGAGGG CCACCTGGTC AGCGGCGCGG CGCTGGGCTT CCTGTCCGGC CGTATCGCGT ACGTCCTCGG TACGGACGGA CCGGCCCTGA 27600
Q  S  P  E  L  E  G  H  L  V  S  G  A  A  L  G  F  L  S  G  R  I  A  Y  V  L  G  T  D  G  P  A  L  T

CCGTGGACAC GGCCTGCTCG TCCTCGCTGG TCGCCCTGCA CCTCGCCGTG CAGGCCCTCC GCAAGGGCGA GTGCGACATG GCCCTCGCCG GTGGTGTCAC 27700
  V  D  T  A  C  S  S  S  L  V  A  L  H  L  A  V  Q  A  L  R  K  G  E  C  D  M  A  L  A  G  G  V  T

                                                                                  XmnI
                                                                                  ▼
GGTCATGCCC AACGCGGACC TGTTCGTGCA GTTCAGCCGG CAGCGCGGGC TGGCCGCGGA CGGCCGGTCG AAGGCGTTCG CCACCTCGGC GGACGGCTTC 27800
 V  M  P  N  A  D  L  F  V  Q  F  S  R  Q  R  G  L  A  A  D  G  R  S  K  A  F  A  T  S  A  D  G  F

                                                                                  BamHI
                                                                                  ▼
GGCCCCGCGG AGGGCGCCGG AGTCCTGCTG GTGGAGCGCC TGTCGGACGC CCGCCGCAAC GGACACCGGA TCCTCGCGGT CGTCCGCGGC AGCGCGGTCA 27900
G  P  A  E  G  A  G  V  L  L  V  E  R  L  S  D  A  R  R  N  G  H  R  I  L  A  V  V  R  G  S  A  V  N

ACCAGGACGG CGCCAGCAAC GGCCTCACGG CTCCGCACGG GCCCTCCCAG CAGCGCGTCA TCCGACGGGC CCTGGCGGAC GCCCGGCTCG CGCCGGGTGA 28000
  Q  D  G  A  S  N  G  L  T  A  P  H  G  P  S  Q  Q  R  V  I  R  R  A  L  A  D  A  R  L  A  P  G  D
```

FIG. 31DDD

pikPKS Sequence

```
      10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CGTGGACGTC GTCGAGGCGC ACGGCACGGG CACGCGGCTC GGCGACCCGA TCGAGGCGCA GGCCCTCATC GCCACCTACG GCCAGGAGAA GAGCAGCGAA 28100
 V  D  V   V  E  A  H   G  T  G   T  R  L   G  D  P  I   E  A  Q   A  L  I   A  T  Y  G   Q  E  K   S  S  E

CAGCCGCTGA GGCTGGGCGC GTTGAAGTCG AACATCGGGC ACACGCAGGC CGCGGCCGGT GTCGCAGGTG TCATCAAGAT GGTCCAGGCG ATGCGCCACG 28200
Q  P  L  R   L  G  A   L  K  S   N  I  G  H   T  Q  A   A  A  G   V  A  G  V   I  K  M   V  Q  A   M  R  H  G

GACTGCTGCC GAAGACGCTG CACGTCGACG AGCCCTCGGA CCAGATCGAC TGGTCGGCGG GCACGGTGGA ACTCCTCACC GAGGCCGTCG ACTGGCCGGA 28300
 L  L  P   K  T  L   H  V  D  E   P  S  D   Q  I  D   W  S  A  G   T  V  E   L  L  T   E  A  V  D   W  P  E

GAAGCAGGAC GGCGGGCTGC GCCGCGCGGC TGTCTCCTCC TTCGGCATCA GCGGGACGAA CGCGCACGTC GTCCTGGAGG AGGCCCCGGC GGTCGAGGAC 28400
 K  Q  D   G  G  L  R   R  A  A   V  S  S   F  G  I  S   G  T  N   A  H  V   V  L  E  E   A  P  A   V  E  D

TCCCCGGCCG TCGAGCCGCC GGCCGGTGGC GGTGTGGTGC CGTGGCCGGT GTCCGCGAAG ACTCCGGCCG CGCTGGACGC CCAGATCGGG CAGCTCGCCG 28500
S  P  A  V   E  P  P   A  G  G   G  V  V  P   W  P  V   S  A  K   T  P  A  A   L  D  A   Q  I  G   Q  L  A  A
```

FIG. 31EEE

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                           BamHI
                             ▼
CGTACGCGGA CGGTCGTACG GACGTGGATC CGGCGGTGGC CGCCCGCGCC CTGGTCGACA GCCGTACGGC GATGGAGCAC CGCGCGGTCG CGGTCGGCGA 28600
  Y  A  D  G  R  T  D  V  D  P  A  V  A  A  R  A  L  V  D  S  R  T  A  M  E  H  R  A  V  A  V  G  D

CAGCCGGGAG GCACTGCGGG ACGCCCTGCG GATGCCGGAA GGACTGGTAC GCGGCACGTC CTCGGACGTG GGCCGGGTGG CGTTCGTCTT CCCCGGCCAG 28700
  S  R  E  A  L  R  D  A  L  R  M  P  E  G  L  V  R  G  T  S  S  D  V  G  R  V  A  F  V  F  P  G  Q

                                                                                BsmI
                                                                                 ▼
GGCACGCAGT GGGCCGGCAT GGGCGCCGAA CTCCTTGACA GCTCACCGGA GTTCGCTGCC TCGATGGCCG AATGCGAGAC CGCGCTCTCC CGCTACGTCG 28800
  G  T  Q  W  A  G  M  G  A  E  L  L  D  S  S  P  E  F  A  A  S  M  A  E  C  E  T  A  L  S  R  Y  V  D

ACTGGTCTCT TGAAGCCGTC GTCCGACAGG AACCCGGCGC ACCCACGCTC GACCGCGTCG ACGTCGTCCA GCCCGTGACC TTCGCTGTCA TGGTCTCGCT 28900
  W  S  L  E  A  V  V  R  Q  E  P  G  A  P  T  L  D  R  V  D  V  V  Q  P  V  T  F  A  V  M  V  S  L

                                                                                ApaLI
                                                                                 ▼
GGCGAAGGTC TGGCAGCACC ACGGCATCAC CCCCCAGGCC GTCGTCGGCC ACTCGCAGGG CGAGATCGCC GCCGCGTACG TCGCCGGTGC ACTCACCCTC 29000
  A  K  V  W  Q  H  H  G  I  T  P  Q  A  V  V  G  H  S  Q  G  E  I  A  A  A  Y  V  A  G  A  L  T  L
```

FIG. 31FFF

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GACGACGCCG CCCGCGTCGT CACCCTGCGC AGCAAGTCCA TCGCCGCCCA CCTCGCCGGC AAGGGCGGCA TGATCTCCCT CGCCCTCGAC GAGGCGGCCG 29100
D  D  A  A   R  V  V   T  L  R   S  K  S  I   A  A  H   L  A  G   K  G  G  M   I  S  L   A  L  D   E  A  A  V

TCCTGAAGCG ACTGAGCGAC TTCGACGGAC TCTCCGTCGC CGCCGTCAAC GGCCCCACCG CCACCGTCGT CTCCGGCGAC CCGACCCAGA TCGAGGAACT 29200
  L  K  R   L  S  D   F  D  G  L   S  V  A   A  V  N   G  P  T  A   T  V  V   S  G  D   P  T  Q  I   E  E  L

CGCCCGCACC TGCGAGGCCG ACGGCGTCCG TGCGCGGATC ATCCCGGTCG ACTACGCCTC CCACAGCCGG CAGGTCGAGA TCATCGAGAA GGAGCTGGCC 29300
 A  R  T   C  E  A  D   G  V  R   A  R  I   I  P  V  D   Y  A  S   H  S  R   Q  V  E  I   I  E  K   E  L  A

                                        PmlI
                                         ▼
GAGGTCCTCG CCGGACTCGC CCCGCAGGCT CCGCACGTGC CGTTCTTCTC CACCCTCGAA GGCACCTGGA TCACCGAGCC GGTGCTCGAC GGCACCTACT 29400
E  V  L  A   G  L  A   P  Q  A   P  H  V  P   F  F  S   T  L  E   G  T  W  I   T  E  P   V  L  D   G  T  Y  W

     KpnI
Acc65I
  ▼  ▼
GGTACCGCAA CCTGCGCCAT CGCGTGGGCT TCGCCCCCGC CGTGGAGACC TTGGCGGTTG ACGGCTTCAC CCACTTCATC GAGGTCAGCG CCCACCCCGT 29500
  Y  R  N   L  R  H   R  V  G  F   A  P  A   V  E  T   L  A  V  D   G  F  T   H  F  I   E  V  S  A   H  P  V
```

FIG. 31GGG

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CCTCACCATG ACCCTCCCCG AGACCGTCAC CGGCCTCGGC ACCCTCCGCC GCGAACAGGG AGGCCAGGAG CGTCTGGTCA CCTCACTCGC CGAAGCCTGG 29600
 L  T  M   T  L  P  E   T  V  T   G  L  G   T  L  R  R   E  Q  G   G  Q  E   R  L  V  T   S  L  A   E  A  W

                                                                                                  Eco47III
                                                                                                      ▼
GCCAACGGCC TCACCATCGA CTGGGCGCCC ATCCTCCCCA CCGCAACCGG CCACCACCCC GAGCTCCCCA CCTACGCCTT CCAGACCGAG CGCTTCTGGC 29700
A  N  G  L   T  I  D   W  A  P   I  L  P  T   A  T  G   H  H  P   E  L  P  T   Y  A  F   Q  T  E   R  F  W  L

     PstI
SfcI
▼  ▼
TGCAGAGCTC CGCGCCCACC AGCGCCGCCG ACGACTGGCG TTACCGCGTC GAGTGGAAGC CGCTGACGGC CTCCGGCCAG GCGGACCTGT CCGGGCGGTG 29800
  Q  S  S   A  P  T   S  A  A  D   D  W  R   Y  R  V   E  W  K  P   L  T  A   S  G  Q   A  D  L  S   G  R  W

GATCGTCGCC GTCGGGAGCG AGCCAGAAGC CGAGCTGCTG GGCGCGCTGA AGGCCGCGGG AGCGGAGGTC GACGTACTGG AAGCCGGGGC GGACGACGAC 29900
 I  V  A   V  G  S  E   P  E  A   E  L  L   G  A  L  K   A  A  G   A  E  V   D  V  L  E   A  G  A   D  D  D

CGTGAGGCCC TCGCCGCCCG GCTCACCGCA CTGACGACCG GCGACGGCTT CACCGGCGTG GTCTCGCTCC TCGACGACCT CGTGCCACAG GTCGCCTGGG 30000
R  E  A  L   A  A  R   L  T  A   L  T  T  G   D  G  F   T  G  V   V  S  L  L   D  D  L   V  P  Q   V  A  W  V
```

FIG. 31HHH

pikPKS Sequence

```
      10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

TGCAGGCACT CGGCGACGCC GGAATCAAGG CGCCCCTGTG GTCCGTCACC CAGGGCGCGG TCTCCGTCGG ACGTCTCGAC ACCCCCGCCG ACCCCGACCG 30100
  Q  A  L   G  D  A   G  I  K  A   P  L  W   S  V  T   Q  G  A  V   S  V  G   R  L  D   T  P  A  D   P  D  R

GGCCATGCTC TGGGGCCTCG GCCGCGTCGT CGCCCTTGAG CACCCCGAAC GCTGGGCCGG CCTCGTCGAC CTCCCCGCCC AGCCCGATGC CGCCGCCCTC 30200
  A  M  L   W  G  L  G   R  V  V   A  L  E   H  P  E  R   W  A  G   L  V  D   L  P  A  Q   P  D  A   A  A  L

                                                    BsaBI
                                                      ▼
GCCCACCTCG TCACCGCACT CTCCGGCGCC ACCGGCGAGG ACCAGATCGC CATCCGCACC ACCGGACTCC ACGCCCGCCG CCTCGCCCGC GCACCCCTCC 30300
A  H  L  V   T  A  L   S  G  A   T  G  E  D   Q  I  A   I  R  T   T  G  L  H   A  R  R   L  A  R   A  P  L  H

ACGGACGTCG GCCCACCCGC GACTGGCAGC CCCACGGCAC CGTCCTCATC ACCGGCGGCA CCGGAGCCCT CGGCAGCCAC GCCGCACGCT GGATGGCCCA 30400
  G  R  R   P  T  R   D  W  Q  P   H  G  T   V  L  I   T  G  G  T   G  A  L   G  S  H   A  A  R  W   M  A  H

CCACGGAGCC GAACACCTCC TCCTCGTCAG CCGCAGCGGC GAACAAGCCC CCGGAGCCAC CCAACTCACC GCCGAACTCA CCGCATCGGG CGCCCGCGTC 30500
  H  G  A   E  H  L  L   L  V  S   R  S  G   E  Q  A  P   G  A  T   Q  L  T   A  E  L  T   A  S  G   A  R  V
```

FIG. 31III

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

ACCATCGCCG CCTGCGACGT CGCCGACCCC CACGCCATGC GCACCCTCCT CGACGCCATC CCCGCCGAGA CGCCCCTCAC CGCCGTCGTC CACACCGCCG 30600
T  I  A  A   C  D  V   A  D  P   H  A  M  R   T  L  L   D  A  I   P  A  E  T   P  L  T   A  V  V   H  T  A  G

GCGCACCGGG CGGCGATCCG CTGGACGTCA CCGGCCCGGA GGACATCGCC CGCATCCTGG GCGCGAAGAC GAGCGGCGCC GAGGTCCTCG ACGACCTGCT 30700
  A  P  G   G  D  P   L  D  V  T   G  P  E   D  I  A   R  I  L  G   A  K  T   S  G  A   E  V  L  D   D  L  L

CCGCGGCACT CCGCTGGACG CCTTCGTCCT CTACTCCTCG AACGCCGGGG TCTGGGGCAG CGGCAGCCAG GGCGTCTACG CGGCGGCCAA CGCCCACCTC 30800
 R  G  T   P  L  D  A   F  V  L   Y  S  S   N  A  G  V   W  G  S   G  S  Q   G  V  Y  A   A  A  N   A  H  L

                                                                                                    MluI
                                                                                                     ▼
GACGCGCTCG CCGCCCGGCG CCGCGCCCGG GGCGAGACGG CGACCTCGGT CGCCTGGGGC CTCTGGGCCG GCGACGGCAT GGGCCGGGGC GCCGACGACG 30900
D  A  L  A   A  R  R   R  A  R   G  E  T  A   T  S  V   A  W  G   L  W  A  G   D  G  M   G  R  G   A  D  D  A

                                                                  MscI
                                                                  BalI
                                                                   ▼
CGTACTGGCA GCGTCGCGGC ATCCGTCCGA TGAGCCCCGA CCGCGCCCTG GACGAACTGG CCAAGGCCCT GAGCCACGAC GAGACCTTCG TCGCCGTGGC 31000
  Y  W  Q   R  R  G   I  R  P  M   S  P  D   R  A  L   D  E  L  A   K  A  L   S  H  D   E  T  F  V   A  V  A
```

FIG. 31JJJ

pikPKS Sequence

```
      10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CGATGTCGAC TGGGAGCGGT TCGCGCCCGC GTTCACGGTG TCCCGTCCCA GCCTTCTGCT CGACGGCGTC CCGGAGGCCC GGCAGGCGCT CGCCGCACCC 31100
 D  V  D   W  E  R  F   A  P  A   F  T  V   S  R  P  S   L  L  L   D  G  V   P  E  A  R   Q  A  L   A  A  P

GTCGGTGCCC CGGCTCCCGG CGACGCCGCC GTGGCGCCGA CCGGGCAGTC GTCGGCGCTG GCCGCGATCA CCGCGCTCCC CGAGCCCGAG CGCCGGCCGG 31200
 V  G  A  P   A  P  G   D  A  A   V  A  P  T   G  Q  S   S  A  L   A  A  I  T   A  L  P   E  P  E   R  R  P  A

CGCTCCTCAC CCTCGTCCGT ACCCACGCGG CGGCCGTACT CGGCCATTCC TCCCCCGACC GGGTGGCCCC CGGCCGTGCC TTCACCGAGC TCGGCTTCGA 31300
 L  L  T   L  V  R   T  H  A  A   A  V  L   G  H  S   S  P  D  R   V  A  P   G  R  A   F  T  E  L   G  F  D

CTCGCTGACG GCCGTGCAGC TCCGCAACCA GCTCTCCACG GTGGTCGGCA ACAGGCTCCC CGCCACCACG GTCTTCGACC ACCCGACGCC CGCCGCACTC 31400
 S  L  T   A  V  Q  L   R  N  Q   L  S  T   V  V  G  N   R  L  P   A  T  T   V  F  D  H   P  T  P   A  A  L

GCCGCGCACC TCCACGAGGC GTACCTCGCA CCGGCCGAGC CGGCCCCGAC GGACTGGGAG GGGCGGGTGC GCCGGGCCCT GGCCGAACTG CCCCTCGACC 31500
 A  A  H  L   H  E  A   Y  L  A   P  A  E  P   A  P  T   D  W  E   G  R  V  R   R  A  L   A  E  L   P  L  D  R
```

FIG. 31KKK

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GGCTGCGGGA CGCGGGGGTC CTCGACACCG TCCTGCGCCT CACCGGCATC GAGCCCGAGC CGGGTTCCGG CGGTTCGGAC GGCGGCGCCG CCGACCCTGG 31600
  L  R  D   A  G  V   L  D  T  V   L  R  L   T  G  I   E  P  E  P   G  S  G   G  S  D   G  G  A  A   D  P  G

TGCGGAGCCG GAGGCGTCGA TCGACGACCT GGACGCCGAG GCCCTGATCC GGATGGCTCT CGGCCCCCGT AACACCTGAC CCGACCGCGG TCCTGCCCCA 31700
 A  E  P   E  A  S  I   D  D  L   D  A  E   A  L  I  R   M  A  L   G  P  R   N  T

CGCGCCGCAC CCCGCGCATC CCGCGCACCA CCCGCCCCCA CACGCCCACA ACCCCATCCA CGAGCGGAAG ACCACACCCA GATGACGAGT TCCAACGAAC 31800
                                                                                M  T  S   S  N  E  Q

AGTTGGTGGA CGCTCTGCGC GCCTCTCTCA AGGAGAACGA AGAACTCCGG AAAGAGAGCC GTCGCCGGGC CGACCGTCGG CAGGAGCCCA TGGCGATCGT 31900
  L  V  D   A  L  R   A  S  L  K   E  N  E   E  L  R   K  E  S  R   R  R  A   D  R  R   Q  E  P  M   A  I  V

                                                                                          KpnI
                                                                                Acc65I
                                                                                  ▼        ▼
CGGCATGAGC TGCCGGTTCG CGGGCGGAAT CCGGTCCCCC GAGGACCTCT GGGACGCCGT CGCCGCGGGC AAGGACCTGG TCTCCGAGGT ACCGGAGGAG 32000
 G  M  S   C  R  F  A   G  G  I   R  S  P   E  D  L  W   D  A  V   A  A  G   K  D  L  V   S  E  V   P  E  E
```

FIG. 31LLL

pikPKS Sequence

```
      10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CGCGGCTGGG ACATCGACTC CCTCTACGAC CCGGTGCCCG GGCGCAAGGG CACGACGTAC GTCCGCAACG CCGCGTTCCT CGACGACGCC GCCGGATTCG 32100
R  G  W  D  I  D  S  L  Y  D  P  V  P  G  R  K  G  T  T  Y  V  R  N  A  A  F  L  D  D  A  A  G  F  D

ACGCGGCCTT CTTCGGGATC TCGCCGCGCG AGGCCCTCGC CATGGACCCG CAGCAGCGGC AGCTCCTCGA AGCCTCCTGG GAGGTCTTCG AGCGGGCCGG 32200
  A  A  F  F  G  I  S  P  R  E  A  L  A  M  D  P  Q  Q  R  Q  L  L  E  A  S  W  E  V  F  E  R  A  G

CATCGACCCC GCGTCGGTCC GCGGCACCGA CGTCGGCGTG TACGTGGGCT GTGGCTACCA GGACTACGCG CCGGACATCC GGGTCGCCCC CGAAGGCACC 32300
 I  D  P  A  S  V  R  G  T  D  V  G  V  Y  V  G  C  G  Y  Q  D  Y  A  P  D  I  R  V  A  P  E  G  T

GGCGGTTACG TCGTCACCGG CAACTCCTCC GCCGTGGCCT CCGGGCGCAT CGCGTACTCC CTCGGCCTGG AGGGACCCGC CGTGACCGTG GACACGGCGT 32400
G  G  Y  V  V  T  G  N  S  S  A  V  A  S  G  R  I  A  Y  S  L  G  L  E  G  P  A  V  T  V  D  T  A  C

GCTCCTCTTC GCTCGTCGCC CTGCACCTCG CCCTGAAGGG CCTGCGGAAC GGCGACTGCT CGACGGCACT CGTGGGCGGC GTGGCCGTCC TCGCGACGCC 32500
  S  S  S  L  V  A  L  H  L  A  L  K  G  L  R  N  G  D  C  S  T  A  L  V  G  G  V  A  V  L  A  T  P
```

FIG. 31MMM

pikPKS Sequence

```
      10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GGGCGCGTTC ATCGAGTTCA GCAGCCAGCA GGCCATGGCC GCCGACGGCC GGACCAAGGG CTTCGCCTCG GCGGCGGACG GCCTCGCCTG GGGCGAGGGC 32600
 G  A  F  I  E  F  S  S  Q  Q  A  M  A  A  D  G  R  T  K  G  F  A  S  A  A  D  G  L  A  W  G  E  G

GTCGCCGTAC TCCTCCTCGA ACGGCTCTCC GACGCGCGGC GCAAGGGCCA CCGGGTCCTG GCCGTCGTGC GCGGCAGCGC CATCAACCAG GACGGCGCGA 32700
 V  A  V  L  L  L  E  R  L  S  D  A  R  R  K  G  H  R  V  L  A  V  V  R  G  S  A  I  N  Q  D  G  A  S

GCAACGGCCT CACGGCTCCG CACGGGCCCT CCCAGCAGCA CCTGATCCGC CAGGCCCTGG CCGACGCGCG GCTCACGTCG AGCGACGTGG ACGTCGTGGA 32800
 N  G  L  T  A  P  H  G  P  S  Q  Q  H  L  I  R  Q  A  L  A  D  A  R  L  T  S  S  D  V  D  V  V  E

                                                                                  AscI
                                                                                   ▼
GGGCCACGGC ACGGGGACCC GTCTCGGCGA CCCGATCGAG GCGCAGGCGC TGCTCGCCAC GTACGGGCAG GGGCGCGCCC CGGGGCAGCC GCTGCGGCTG 32900
 G  H  G  T  G  T  R  L  G  D  P  I  E  A  Q  A  L  L  A  T  Y  G  Q  G  R  A  P  G  Q  P  L  R  L

GGGACGCTGA AGTCGAACAT CGGGCACACG CAGGCCGCTT CGGGTGTCGC CGGTGTCATC AAGATGGTGC AGGCGCTGCG CCACGGGGTG CTGCCGAAGA 33000
 G  T  L  K  S  N  I  G  H  T  Q  A  A  S  G  V  A  G  V  I  K  M  V  Q  A  L  R  H  G  V  L  P  K  T
```

FIG. 31NNN

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

       PmlI
       ▼
CCCTGCACGT GGACGAGCCG ACGGACCAGG TCGACTGGTC GGCCGGTTCG GTCGAGCTGC TCACCGAGGC CGTGGACTGG CCGGAGCGGC CGGGCCGGCT 33100
  L  H  V   D  E  P   T  D  Q  V   D  W  S   A  G  S   V  E  L  L   T  E  A   V  D  W   P  E  R  P   G  R  L

CCGCCGGGCG GGCGTCTCCG CGTTCGGCGT GGGCGGGACG AACGCGCACG TCGTCCTGGA GGAGGCCCCG GCGGTCGAGG AGTCCCCTGC CGTCGAGCCG 33200
  R  R  A   G  V  S  A   F  G  V   G  G  T   N  A  H  V   V  L  E   E  A  P   A  V  E  E   S  P  A   V  E  P

CCGGCCGGTG GCGGCGTGGT GCCGTGGCCG GTGTCCGCGA AGACCTCGGC CGCACTGGAC GCCCAGATCG GGCAGCTCGC CGCATACGCG GAAGACCGCA 33300
P  A  G  G   G  V  V   P  W  P   V  S  A  K   T  S  A   A  L  D   A  Q  I  G   Q  L  A   A  Y  A   E  D  R  T

       BamHI
       ▼
CGGACGTGGA TCCGGCGGTG GCCGCCCGCG CCCTGGTCGA CAGCCGTACG GCGATGGAGC ACCGCGCGGT CGCGGTCGGC GACAGCCGGG AGGCACTGCG 33400
  D  V  D   P  A  V   A  A  R  A   L  V  D   S  R  T   A  M  E  H   R  A  V   A  V  G   D  S  R  E   A  L  R

GGACGCCCTG CGGATGCCGG AAGGACTGGT ACGGGGCACG GTCACCGATC CGGGCCGGGT GGCGTTCGTC TTCCCCGGCC AGGGCACGCA GTGGGCCGGC 33500
  D  A  L   R  M  P  E   G  L  V   R  G  T   V  T  D  P   G  R  V   A  F  V   F  P  G  Q   G  T  Q   W  A  G
```

FIG. 31000

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                  EcoRI
                                  ApoI                                    BsmI
                                    ▼                                       ▼
ATGGGCGCCG AACTCCTCGA CAGCTCACCC GAATTCGCCG CCGCCATGGC CGAATGCGAG ACCGCACTCT CCCCGTACGT CGACTGGTCT CTCGAAGCCG 33600
M  G  A  E  L  L  D  S  S  P  E  F  A  A  A  M  A  E  C  E  T  A  L  S  P  Y  V  D  W  S  L  E  A  V

TCGTCCGACA GGCTCCCAGC GCACCGACAC TCGACCGCGT CGACGTCGTC CAGCCCGTCA CCTTCGCCGT CATGGTCTCC CTCGCCAAGG TCTGGCAGCA 33700
  V  R  Q  A  P  S  A  P  T  L  D  R  V  D  V  V  Q  P  V  T  F  A  V  M  V  S  L  A  K  V  W  Q  H

CCACGGCATC ACCCCCGAGG CCGTCATCGG CCACTCCCAG GGCGAGATCG CCGCCGCGTA CGTCGCCGGT GCCCTCACCC TCGACGACGC CGCTCGTGTC 33800
 H  G  I  T  P  E  A  V  I  G  H  S  Q  G  E  I  A  A  A  Y  V  A  G  A  L  T  L  D  D  A  A  R  V

GTGACCCTCC GCAGCAAGTC CATCGCCGCC CACCTCGCCG GCAAGGGCGG CATGATCTCC CTCGCCCTCA GCGAGGAAGC CACCCGGCAG CGCATCGAGA 33900
V  T  L  R  S  K  S  I  A  A  H  L  A  G  K  G  G  M  I  S  L  A  L  S  E  E  A  T  R  Q  R  I  E  N

ACCTCCACGG ACTGTCGATC GCCGCCGTCA ACGGGCCTAC CGCCACCGTG GTTTCGGGCG ACCCCACCCA GATCCAAGAA CTTGCTCAGG CGTGTGAGGC 34000
  L  H  G  L  S  I  A  A  V  N  G  P  T  A  T  V  V  S  G  D  P  T  Q  I  Q  E  L  A  Q  A  C  E  A
```

FIG. 31PPP

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CGACGGCATC CGCGCACGGA TCATCCCCGT CGACTACGCC TCCCACAGCG CCCACGTCGA GACCATCGAG AACGAACTCG CCGACGTCCT GGCGGGGTTG 34100
 D  G  I   R  A  R  I   I  P  V   D  Y  A   S  H  S  A   H  V  E   T  I  E   N  E  L  A   D  V  L   A  G  L

                                                                                    KpnI
                                                                                  Acc65I
                                                                                    ▼  ▼
TCCCCCCAGA CACCCCAGGT CCCCTTCTTC TCCACCCTCG AAGGCACCTG GATCACCGAA CCCGCCCTCG ACGGCGGCTA CTGGTACCGC AACCTCCGCC 34200
 S  P  Q  T   P  Q  V   P  F  F   S  T  L  E   G  T  W   I  T  E   P  A  L  D   G  G  Y   W  Y  R   N  L  R  H

ATCGTGTGGG CTTCGCCCCG GCCGTCGAGA CCCTCGCCAC CGACGAAGGC TTCACCCACT TCATCGAGGT CAGCGCCCAC CCCGTCCTCA CCATGACCCT 34300
  R  V  G   F  A  P   A  V  E  T   L  A  T   D  E  G   F  T  H  F   I  E  V   S  A  H   P  V  L  T   M  T  L

                        MscI
                        BalI
                          ▼
CCCCGACAAG GTCACCGGCC TGGCCACCCT CCGACGCGAG GACGGCGGAC AGCACCGCCT CACCACCTCC CTTGCCGAGG CCTGGGCCAA CGGCCTCGCC 34400
 P  D  K   V  T  G  L   A  T  L   R  R  E   D  G  G  Q   H  R  L   T  T  S   L  A  E  A   W  A  N   G  L  A

CTCGACTGGG CCTCCCTCCT GCCCGCCACG GGCGCCCTCA GCCCCGCCGT CCCCGACCTC CCGACGTACG CCTTCCAGCA CCGCTCGTAC TGGATCAGCC 34500
 L  D  W  A   S  L  L   P  A  T   G  A  L  S   P  A  V   P  D  L   P  T  Y  A   F  Q  H   R  S  Y   W  I  S  P
```

FIG. 31QQQ

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CCGCGGGTCC CGGCGAGGCG CCCGCGCACA CCGCTTCCGG GCGCGAGGCC GTCGCCGAGA CGGGGCTCGC GTGGGGCCCG GGTGCCGAGG ACCTCGACGA 34600
  A  G  P   G  E  A   P  A  H  T   A  S  G   R  E  A   V  A  E  T   G  L  A   W  G  P   G  A  E  D   L  D  E

GGAGGGCCGG CGCAGCGCCG TACTCGCGAT GGTGATGCGG CAGGCGGCCT CCGTGCTCCG GTGCGACTCG CCCGAAGAGG TCCCCGTCGA CCGCCCGCTG 34700
  E  G  R   R  S  A  V   L  A  M   V  M  R   Q  A  A  S   V  L  R   C  D  S   P  E  E  V   P  V  D   R  P  L

CGGGAGATCG GCTTCGACTC GCTGACCGCC GTCGACTTCC GCAACCGCGT CAACCGGCTG ACCGGTCTCC AGCTGCCGCC CACCGTCGTG TTCCAGCACC 34800
  R  E  I  G   F  D  S   L  T  A   V  D  F  R   N  R  V   N  R  L   T  G  L  Q   L  P  P   T  V  V   F  Q  H  P
                  *                                                                                        ACP 6

CGACGCCCGT CGCGCTCGCC GAGCGCATCA GCGACGAGCT GGCCGAGCGG AACTGGGCCG TCGCCGAGCC GTCGGATCAC GAGCAGGCGG AGGAGGAGAA 34900
  T  P  V   A  L  A   E  R  I  S   D  E  L   A  E  R   N  W  A  V   A  E  P   S  D  H   E  Q  A  E   E  E  K
                                              →                                                          TE domain

GGCCGCCGCT CCGGCGGGGG CCCGCTCCGG GGCCGACACC GGCGCCGGCG CCGGGATGTT CCGCGCCCTG TTCCGGCAGG CCGTGGAGGA CGACCGGTAC 35000
  A  A  A   P  A  G  A   R  S  G   A  D  T   G  A  G  A   G  M  F   R  A  L   F  R  Q  A   V  E  D   D  R  Y
```

FIG. 31RRR

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

GGCGAGTTCC TCGACGTCCT CGCCGAAGCC TCCGCGTTCC GCCCGCAGTT CGCCTCGCCC GAGGCCTGCT CGGAGCGGCT CGACCCGGTG CTGCTCGCCG 35100
G  E  F  L  D  V  L  A  E  A  S  A  F  R  P  Q  F  A  S  P  E  A  C  S  E  R  L  D  P  V  L  L  A  G

                                                                                    EspI
                                                                                    Bpu1102I
                                                                                    ▼
GCGGTCCGAC GGACCGGGCG GAAGGCCGTG CCGTTCTCGT CGGCTGCACC GGCACCGCGG CGAACGGCGG CCCGCACGAG TTCCTGCGGC TCAGCACCTC 35200
  G  P  T  D  R  A  E  G  R  A  V  L  V  G  C  T  G  T  A  A  N  G  G  P  H  E  F  L  R  L  S  T  S

CTTCCAGGAG GAGCGGGACT TCCTCGCCGT ACCTCTCCCC GGCTACGGCA CGGGTACGGG CACCGGCACG GCCCTCCTCC CGGCCGATCT CGACACCGCG 35300
 F  Q  E  E  R  D  F  L  A  V  P  L  P  G  Y  G  T  G  T  G  T  G  T  A  L  L  P  A  D  L  D  T  A

CTCGACGCCC AGGCCCGGGC GATCCTCCGG GCCGCCGGGG ACGCCCCGGT CGTCCTGCTC GGGCACTCCG GCGGCGCCCT GCTCGCGCAC GAGCTGGCCT 35400
L  D  A  Q  A  R  A  I  L  R  A  A  G  D  A  P  V  V  L  L  G  H  S  G  G  A  L  L  A  H  E  L  A  F
                                                                  *

                      AscI
                      ▼
TCCGCCTGGA GCGGGCGCAC GGCGCGCCGC CGGCCGGGAT CGTCCTGGTC GACCCCTATC CGCCGGGCCA TCAGGAGCCC ATCGAGGTGT GGAGCAGGCA 35500
  R  L  E  R  A  H  G  A  P  P  A  G  I  V  L  V  D  P  Y  P  P  G  H  Q  E  P  I  E  V  W  S  R  Q
```

FIG. 31SSS

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                                    MscI
                                                    BalI
                                                     ▼
GCTGGGCGAG GGCCTGTTCG CGGGCGAGCT GGAGCCGATG TCCGATGCGC GGCTGCTGGC CATGGGCCGG TACGCGCGGT TCCTCGCCGG CCCGCGGCCG 35600
 L  G  E  G  L  F  A  G  E  L  E  P  M  S  D  A  R  L  L  A  M  G  R  Y  A  R  F  L  A  G  P  R  P

GGCCGCAGCA GCGCGCCCGT GCTTCTGGTC CGTGCCTCCG AACCGCTGGG CGACTGGCAG GAGGAGCGGG GCGACTGGCG TGCCCACTGG GACCTTCCGC 35700
G  R  S  S  A  P  V  L  L  V  R  A  S  E  P  L  G  D  W  Q  E  E  R  G  D  W  R  A  H  W  D  L  P  H

ACACCGTCGC GGACGTGCCG GGCGACCACT TCACGATGAT GCGGGACCAC GCGCCGGCCG TCGCCGAGGC CGTCCTCTCC TGGCTCGACG CCATCGAGGG 35800
 T  V  A  D  V  P  G  D  H  F  T  M  M  R  D  H  A  P  A  V  A  E  A  V  L  S  W  L  D  A  I  E  G
                          *

                                                    BamHI
                                                     ▼
CATCGAGGGG GCGGGCAAGT GACCGACAGA CCTCTGAACG TGGACAGCGG ACTGTGGATC CGGCGCTTCC ACCCCGCGCC GAACAGCGCG GTGCGGCTGG 35900
 I  E  G  A  G  K
                V  T  D  R  P  L  N  V  D  S  G  L  W  I  R  R  F  H  P  A  P  N  S  A  V  R  L  V
                →                                                                              TE2

TCTGCCTGCC GCACGCCGGC GGCTCCGCCA GCTACTTCTT CCGCTTCTCG GAGGAGCTGC ACCCCTCCGT CGAGGCCCTG TCGGTGCAGT ATCCGGGCCG 36000
 C  L  P  H  A  G  G  S  A  S  Y  F  F  R  F  S  E  E  L  H  P  S  V  E  A  L  S  V  Q  Y  P  G  R
```

FIG. 31TTT

pikPKS Sequence

```
        10         20         30         40         50         60         70         80         90         100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

CCAGGACCGG CGTGCCGAGC CGTGTCTGGA GAGCGTCGAG GAGCTCGCCG AGCATGTGGT CGCGGCCACC GAACCCTGGT GGCAGGAGGG CCGGCTGGCC 36100

 Q  D  R  R  A  E  P  C  L  E  S  V  E  E  L  A  E  H  V  V  A  A  T  E  P  W  W  Q  E  G  R  L  A

TTCTTCGGGC ACAGCCTCGG CGCCTCCGTC GCCTTCGAGA CGGCCCGCAT CCTGGAACAG CGGCACGGGG TACGGCCCGA GGGCCTGTAC GTCTCCGGTC 36200

 F  F  G  H  S  L  G  A  S  V  A  F  E  T  A  R  I  L  E  Q  R  H  G  V  R  P  E  G  L  Y  V  S  G  R
          *

                                                                                  EspI
                                                                                  Bpu1102I
   AscI
    ▼                                                                             ▼
GGCGCGCCCC GTCGCTGGCG CCGGACCGGC TCGTCCACCA GCTGGACGAC CGGGCGTTCC TGGCCGAGAT CCGGCGGCTC AGCGGCACCG ACGAGCGGTT 36300

  R  A  P  S  L  A  P  D  R  L  V  H  Q  L  D  D  R  A  F  L  A  E  I  R  R  L  S  G  T  D  E  R  F

CCTCCAGGAC GACGAGCTGC TGCGGCTGGT GCTGCCCGCG CTGCGCAGCG ACTACAAGGC GGCGGAGACG TACCTGCACC GGCCGTCCGC CAAGCTCACC 36400

 L  Q  D  D  E  L  L  R  L  V  L  P  A  L  R  S  D  Y  K  A  A  E  T  Y  L  H  R  P  S  A  K  L  T

TGCCCGGTGA TGGCCCTGGC CGGCGACCGT GACCCGAAGG CGCCGCTGAA CGAGGTGGCC GAGTGGCGTC GGCACACCAG CGGGCCGTTC TGCCTCCGGG 36500

 C  P  V  M  A  L  A  G  D  R  D  P  K  A  P  L  N  E  V  A  E  W  R  R  H  T  S  G  P  F  C  L  R  A
```

FIG. 31UUU

pikPKS Sequence

```
      10        20        30        40        50        60        70        80        90        100
1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890 1234567890

                                          BglII                                        AscI
                                            ▼                                          ▼
CGTACTCCGG CGGCCACTTC TACCTCAACG ACCAGTGGCA CGAGATCTGC AACGACATCT CCGACCACCT GCTCGTCACC CGCGGCGCGC CCGATGCCCG 36600

 Y  S  G   G  H  F  Y   L  N  D   Q  W  H   E  I  C  N   D  I  S   D  H  L   L  V  T  R   G  A  P   D  A  R
              *

CGTCGTGCAG CCCCCGACCA GCCTTATCGA AGGAGCGGCG AGAAGATGGC AGAACCCACG GTGACCGACG ACCTGACGGG GGCCCTCACG CAGCCCCCGC 36700

 V  V  Q   P  P  T  S   L  I  E   G  A  A   K  R  W  Q   N  P  R

TGGGCCGCAC CGTCCGCGCG GTGGCCGACC GTGAACTCGG CACCCACCTC CTGGAGACCC GCGGCATCCA CTGGATCC                      36778
```

FIG. 31VV

sugar.finalgene b-1 Sequence

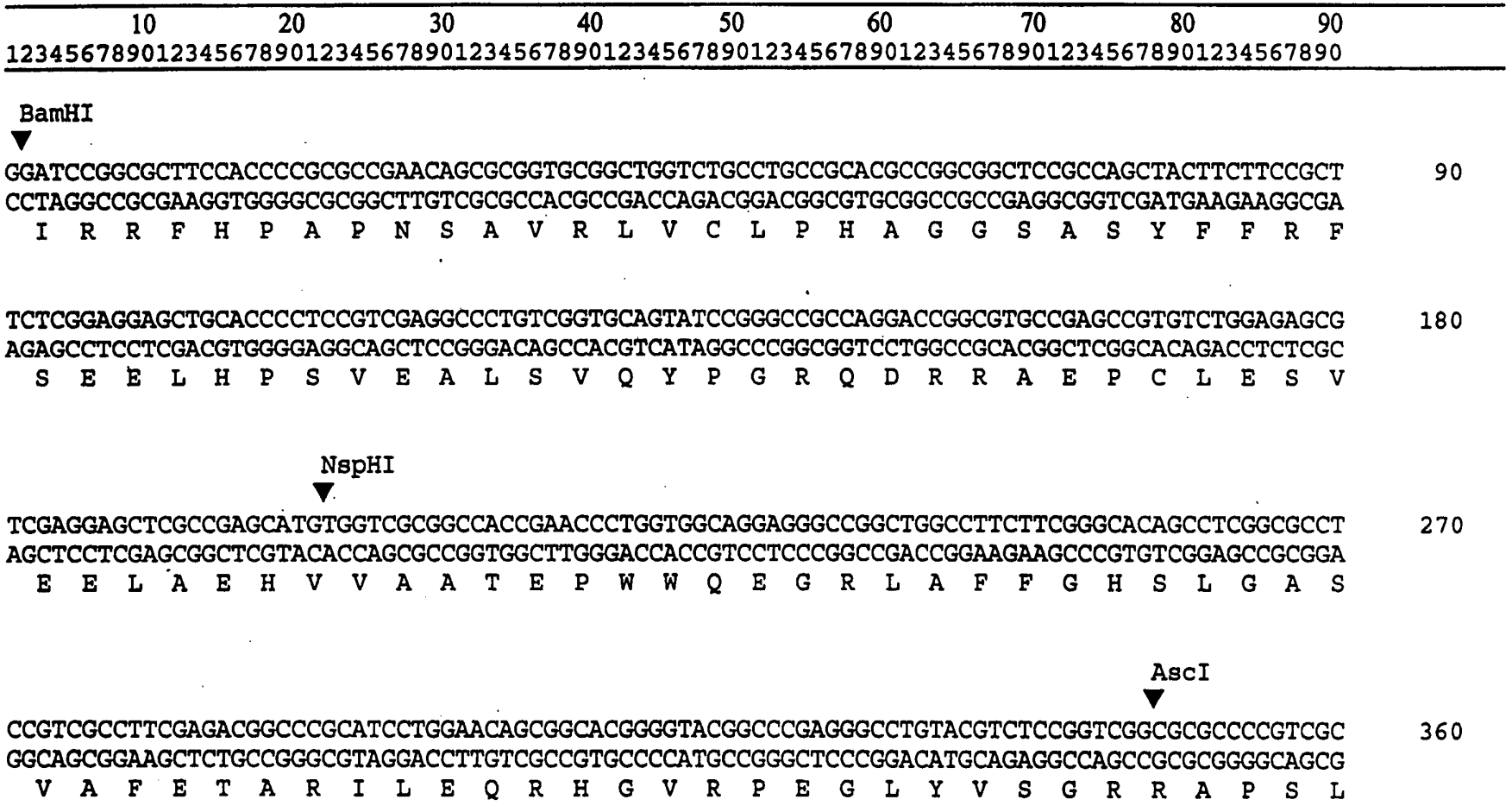

FIG. 32A

sugar.finalgene b-1 Sequence

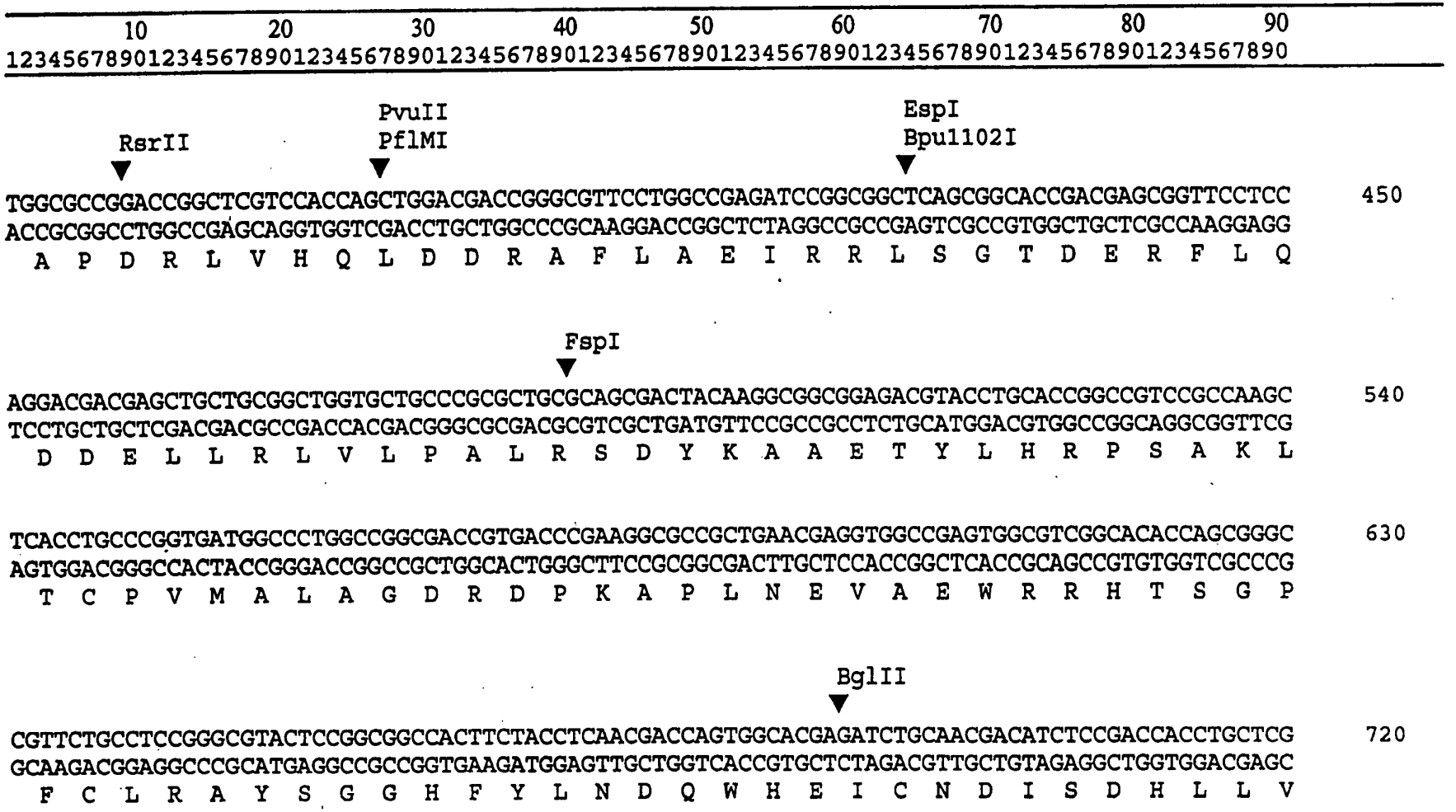

FIG. 32B

sugar.finalgene b-1 Sequence

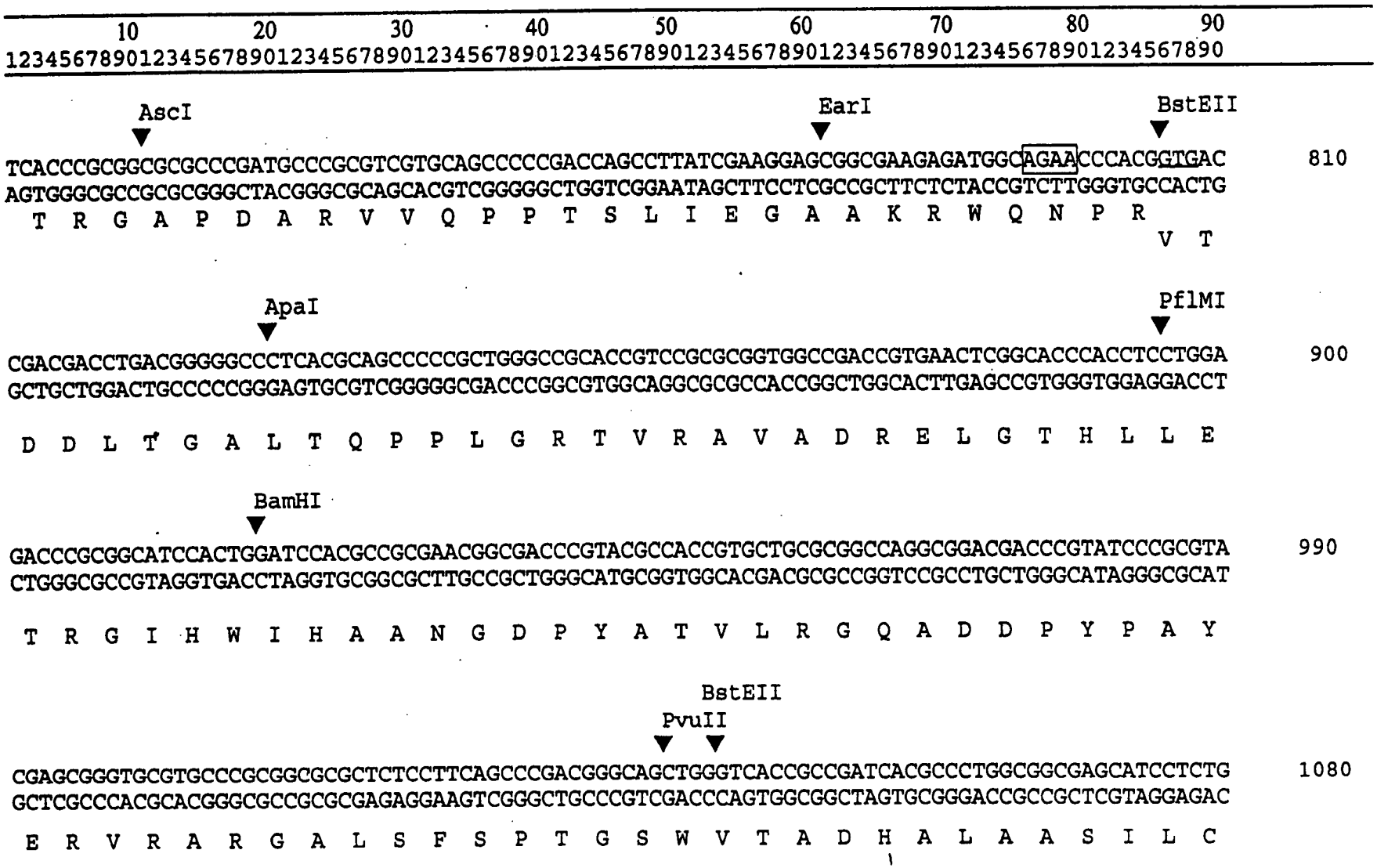

FIG. 32C

sugar.finalgene b-1 Sequence

```
         10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
CTCGACGGACTTCGGGGTCTCCGGCGCCGACGGCGTCCCGGTGCCGCAGCAGGTCCTCTCGTACGGGGAGGGCTGTCCGCTGGAGCGCGA      1170
GAGCTGCCTGAAGCCCCAGAGGCCGCGGCTGCCGCAGGGCCACGGCGTCGTCCAGGAGAGCATGCCCCTCCCGACAGGCGACCTCGCGCT

 S  T  D  F  G  V  S  G  A  D  G  V  P  V  P  Q  Q  V  L  S  Y  G  E  G  C  P  L  E  R  E

       AlwNI                                                    BamHI
       ▼                                                        ▼
GCAGGTGCTGCCGGCGGCCGGTGACGTGCCGGAGGGCGGGCAGCGTGCCGTGGTCGAGGGGATCCACCGGGAGACGCTGGAGGGTCTCGC      1260
CGTCCACGACGGCCGCCGGCCACTGCACGGCCTCCCGCCCGTCGCACGGCACCAGCTCCCCTAGGTGGCCCTCTGCGACCTCCCAGAGCG

 Q  V  L  P  A  A  G  D  V  P  E  G  G  Q  R  A  V  V  E  G  I  H  R  E  T  L  E  G  L  A

GCCGGACCCGTCGGCGTCGTACGCCTTCGAGCTGCTGGGCGGTTTCGTCCGCCCGGCGGTGACGGCCGCTGCCGCCGCCGTGCTGGGTGT      1350
CGGCCTGGGCAGCCGCAGCATGCGGAAGCTCGACGACCCGCCAAAGCAGGCGGGCCGCCACTGCCGGCGACGGCGGCGGCACGACCCACA

 P  D  P  S  A  S  Y  A  F  E  L  L  G  G  F  V  R  P  A  V  T  A  A  A  A  A  V  L  G  V

       RsrII                                                                    AlwNI
       ▼                                                                        ▼
TCCCGCGGACCGGCGCGCGGACTTCGCGGATCTGCTGGAGCGGCTCCGGCCGCTGTCCGACAGCCTGCTGGCCCCGCAGTCCCTGCGGAC      1440
AGGGCGCCTGGCCGCGCGCCTGAAGCGCCTAGACGACCTCGCCGAGGCCGGCGACAGGCTGTCGGACGACCGGGGCGTCAGGGACGCCTG

 P  A  D  R  R  A  D  F  A  D  L  L  E  R  L  R  P  L  S  D  S  L  L  A  P  Q  S  L  R  T
```

FIG. 32D

sugar.finalgene b-1 Sequence

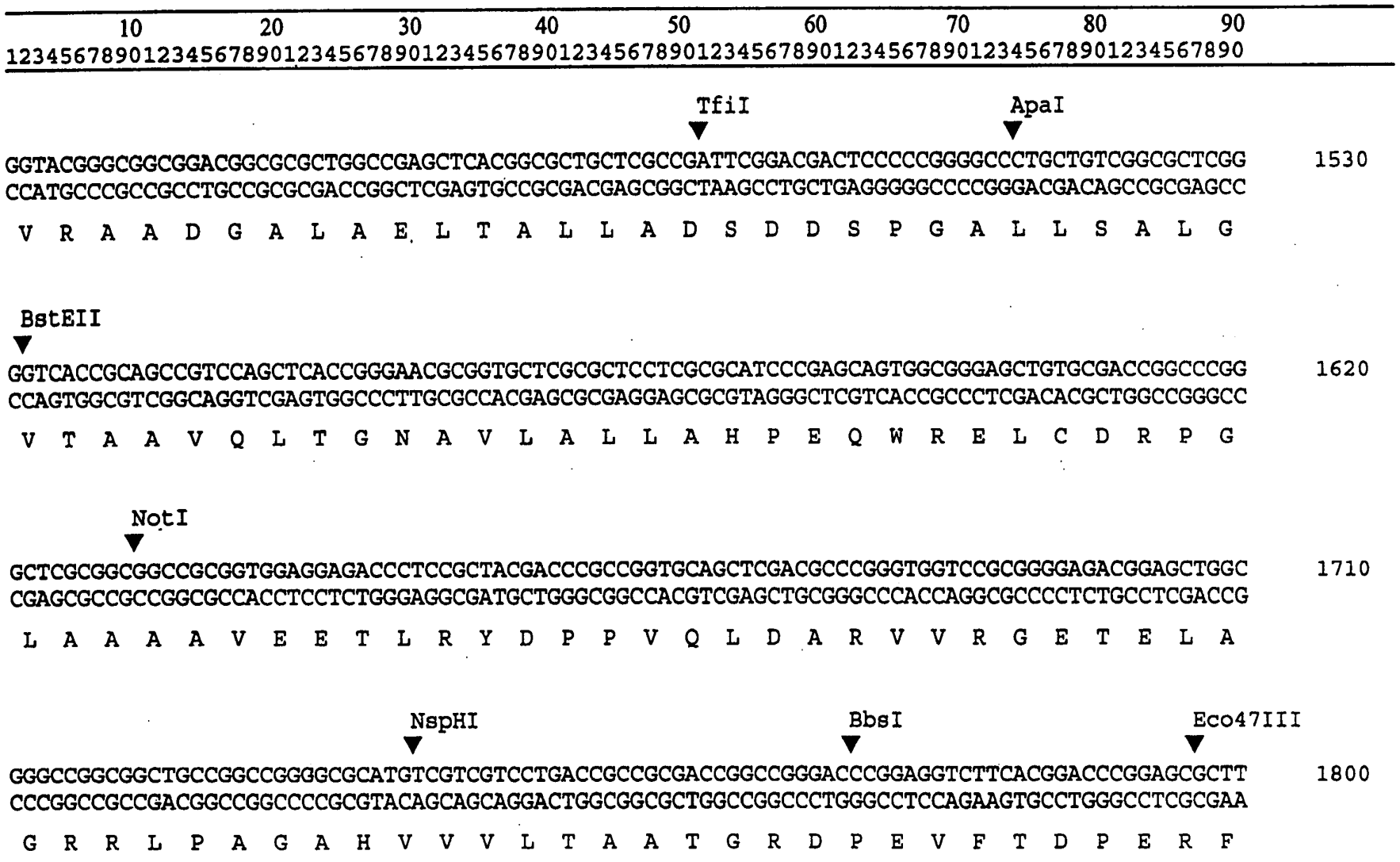

FIG. 32E

sugar.finalgene b-1 Sequence

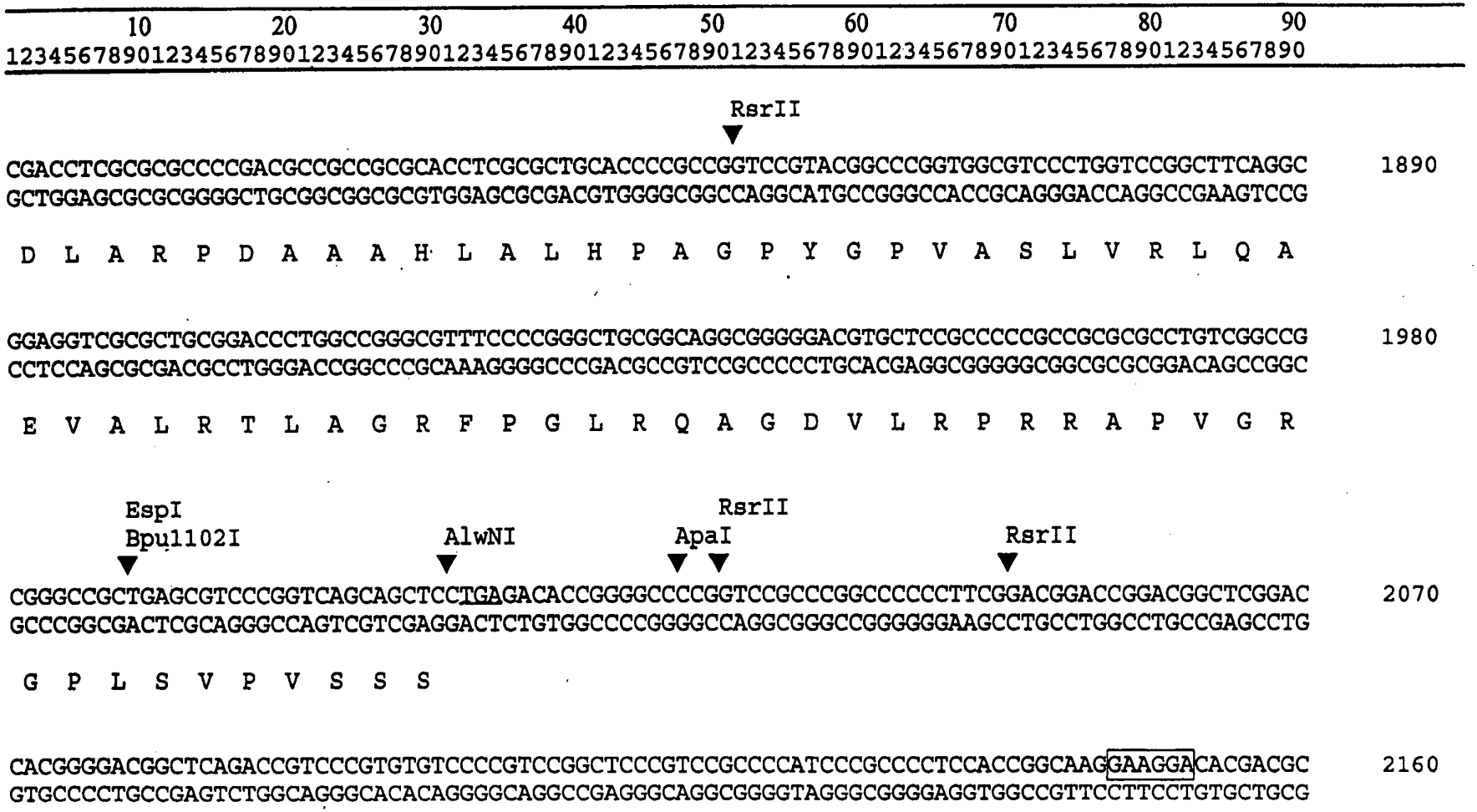

FIG. 32F

sugar.finalgene b-1 Sequence

```
        10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
CATGCGCGTCCTGCTGACCTCGTTCGCACATCACACGCACTACTACGGCCTGGTGCCCCTGGCCTGGGCGCTGCTCGCCGCCGGGCACGA      2250
GTACGCGCAGGACGACTGGAGCAAGCGTGTAGTGTGCGTGATGATGCCGGACCACGGGGACCGGACCCGCGACGAGCGGCGGCCCGTGCT
 M  R  V  L  L  T  S  F  A  H  H  T  H  Y  Y  G  L  V  P  L  A  W  A  L  L  A  A  G  H  E

 DraIII
 ▼
GGTGCGGGTCGCCAGCCAGCCCGCGCTCACGGACACCATCACCGGGTCCGGGCTCGCCGCGGTGCCGGTCGGCACCGACCACCTCATCCA      2340
CCACGCCCAGCGGTCGGTCGGGCGCGAGTGCCTGTGGTAGTGGCCCAGGCCCGAGCGGCGCCACGGCCAGCCGTGGCTGGTGGAGTAGGT
 V  R  V  A  S  Q  P  A  L  T  D  T  I  T  G  S  G  L  A  A  V  P  V  G  T  D  H  L  I  H

                                                  PvuI
                                                  ▼
CGAGTACCGGGTGCGGATGGCGGGCGAGCCGCGCCCGAACCATCCGGCGATCGCCTTCGACGAGGCCCGTCCCGAGCCGCTGGACTGGGA      2430
GCTCATGGCCCACGCCTACCGCCCGCTCGGCGCGGGCTTGGTAGGCCGCTAGCGGAAGCTGCTCCGGGCAGGGCTCGGCGACCTGACCCT
 E  Y  R  V  R  M  A  G  E  P  R  P  N  H  P  A  I  A  F  D  E  A  R  P  E  P  L  D  W  D

CCACGCCCTCGGCATCGAGGCGATCCTCGCCCCGTACTTCCATCTGCTCGCCAACAACGACTCGATGGTCGACGACCTCGTCGACTTCGC      2520
GGTGCGGGAGCCGTAGCTCCGCTAGGAGCGGGGCATGAAGGTAGACGAGCGGTTGTTGCTGAGCTACCAGCTGCTGGAGCAGCTGAAGCG
 H  A  L  G  I  E  A  I  L  A  P  Y  F  H  L  L  A  N  N  D  S  M  V  D  D  L  V  D  F  A
```

FIG. 32G

sugar.finalgene b-1 Sequence

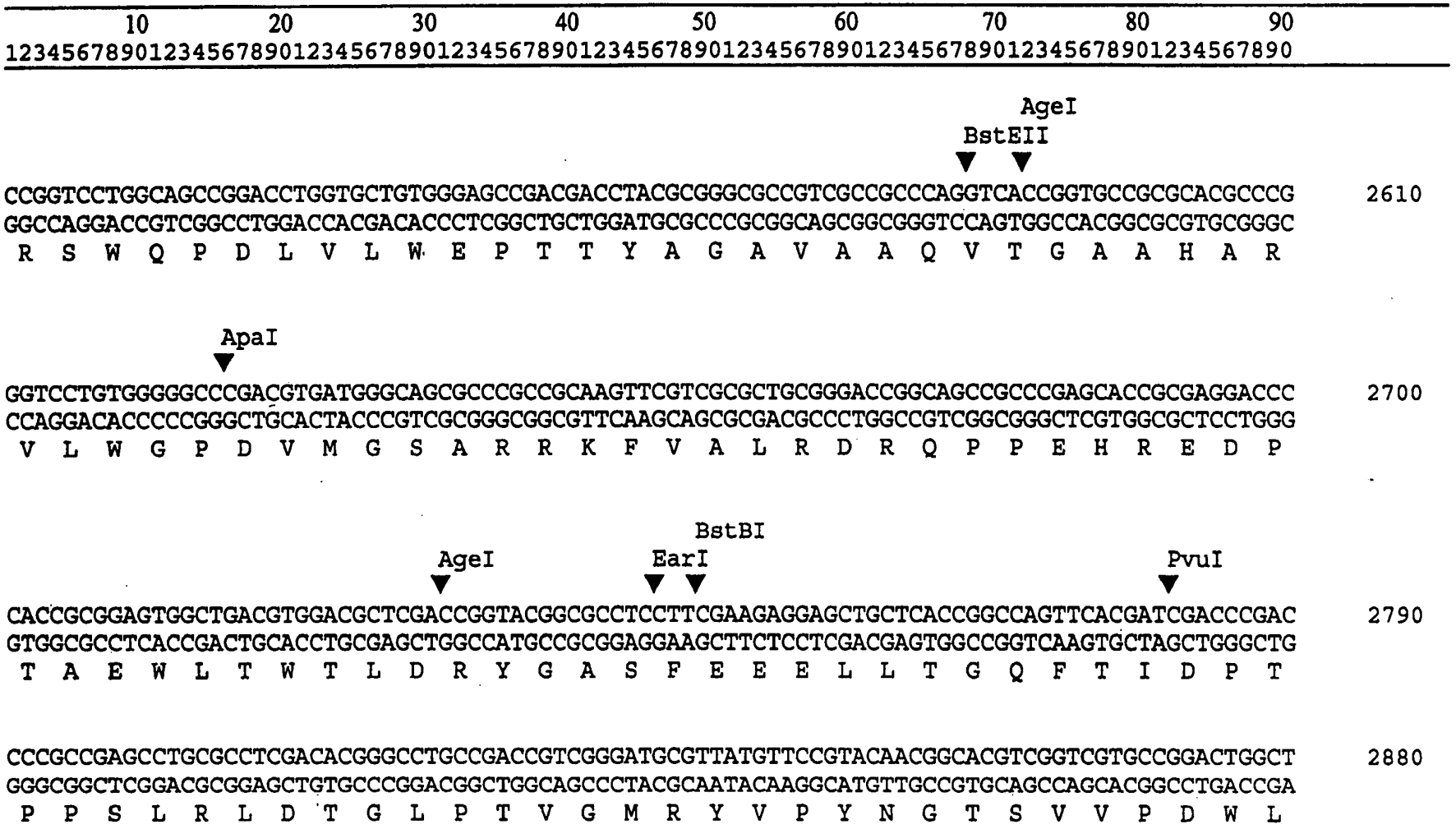

FIG. 32H

sugar.finalgene b-1 Sequence

```
         10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
GAGTGAGCCGCCCGCGCGGCCCCGGGTCTGCCTGACCCTCGGCGTCTCCGCGCGTGAGGTCCTCGGCGGCGACGGCGTCTCGCAGGGCGA      2970
CTCACTCGGCGGGCGCGCCGGGGCCCAGACGGACTGGGAGCCGCAGAGGCGCGCACTCCAGGAGCCGCCGCTGCCGCAGAGCGTCCCGCT
 S  E  P  P  A  R  P  R  V  C  L  T  L  G  V  S  A  R  E  V  L  G  G  D  G  V  S  Q  G  D

CATCCTGGAGGCGCTCGCCGACCTCGACATCGAGCTCGTCGCCACGCTCGACGCGAGTCAGCGCGCCGAGATCCGCAACTACCCGAAGCA      3060
GTAGGACCTCCGCGAGCGGCTGGAGCTGTAGCTCGAGCAGCGGTGCGAGCTGCGCTCAGTCGCGCGGCTCTAGGCGTTGATGGGCTTCGT
 I  L  E  A  L  A  D  L  D  I  E  L  V  A  T  L  D  A  S  Q  R  A  E  I  R  N  Y  P  K  H

CACCCGGTTCACGGACTTCGTGCCGATGCACGCGCTCCTGCCGAGCTGCTCGGCGATCATCCACCACGGCGGGGCGGGCACCTACGCGAC      3150
GTGGGCCAAGTGCCTGAAGCACGGCTACGTGCGCGAGGACGGCTCGACGAGCCGCTAGTAGGTGGTGCCGCCCCGCCCGTGGATGCGCTG
 T  R  F  T  D  F  V  P  M  H  A  L  L  P  S  C  S  A  I  I  H  H  G  G  A  G  T  Y  A  T

      BclI
      ▼
CGCCGTGATCAACGCGGTGCCGCAGGTCATGCTCGCCGAGCTGTGGGACGCGCCGGTCAAGGCGCGGGCCGTCGCCGAGCAGGGGGCGGG      3240
GCGGCACTAGTTGCGCCACGGCGTCCAGTACGAGCGGCTCGACACCCTGCGCGGCCAGTTCCGCGCCCGGCAGCGGCTCGTCCCCCGCCC
 A  V  I  N  A  V  P  Q  V  M  L  A  E  L  W  D  A  P  V  K  A  R  A  V  A  E  Q  G  A  G
```

FIG. 321

sugar.finalgene b-1 Sequence

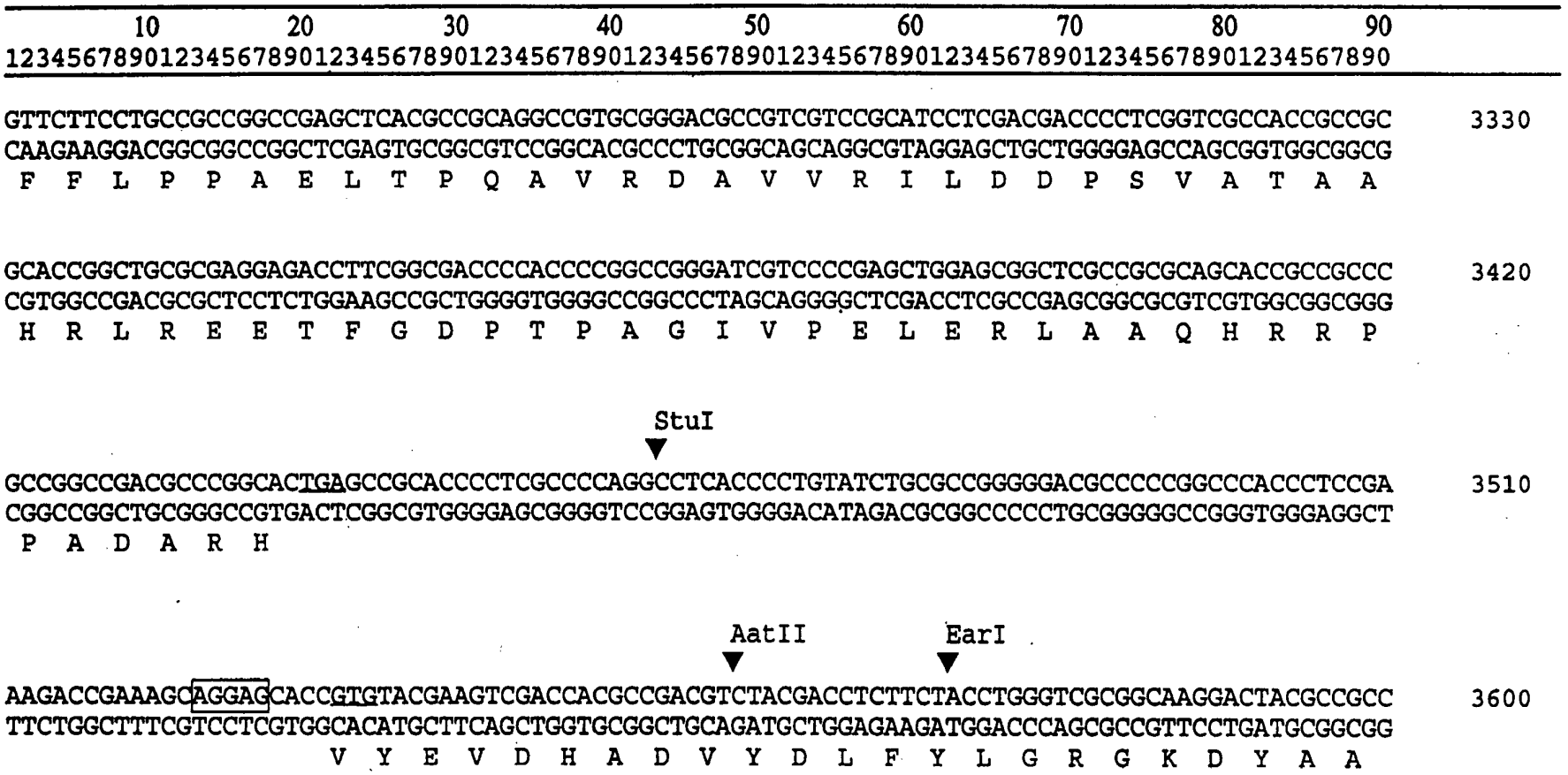

FIG. 32J

sugar.finalgene b-1 Sequence

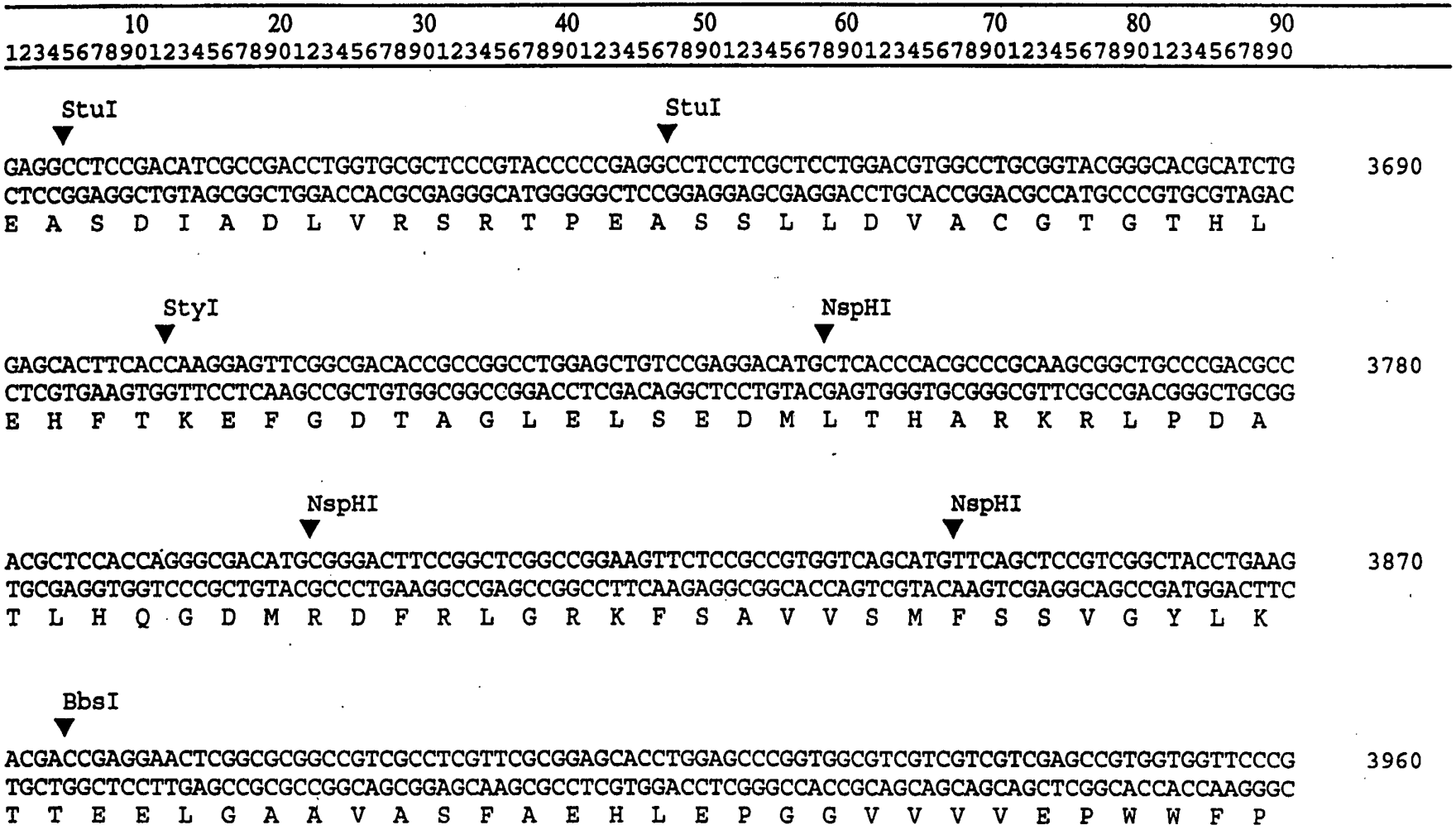

FIG. 32K

sugar.finalgene b-1 Sequence

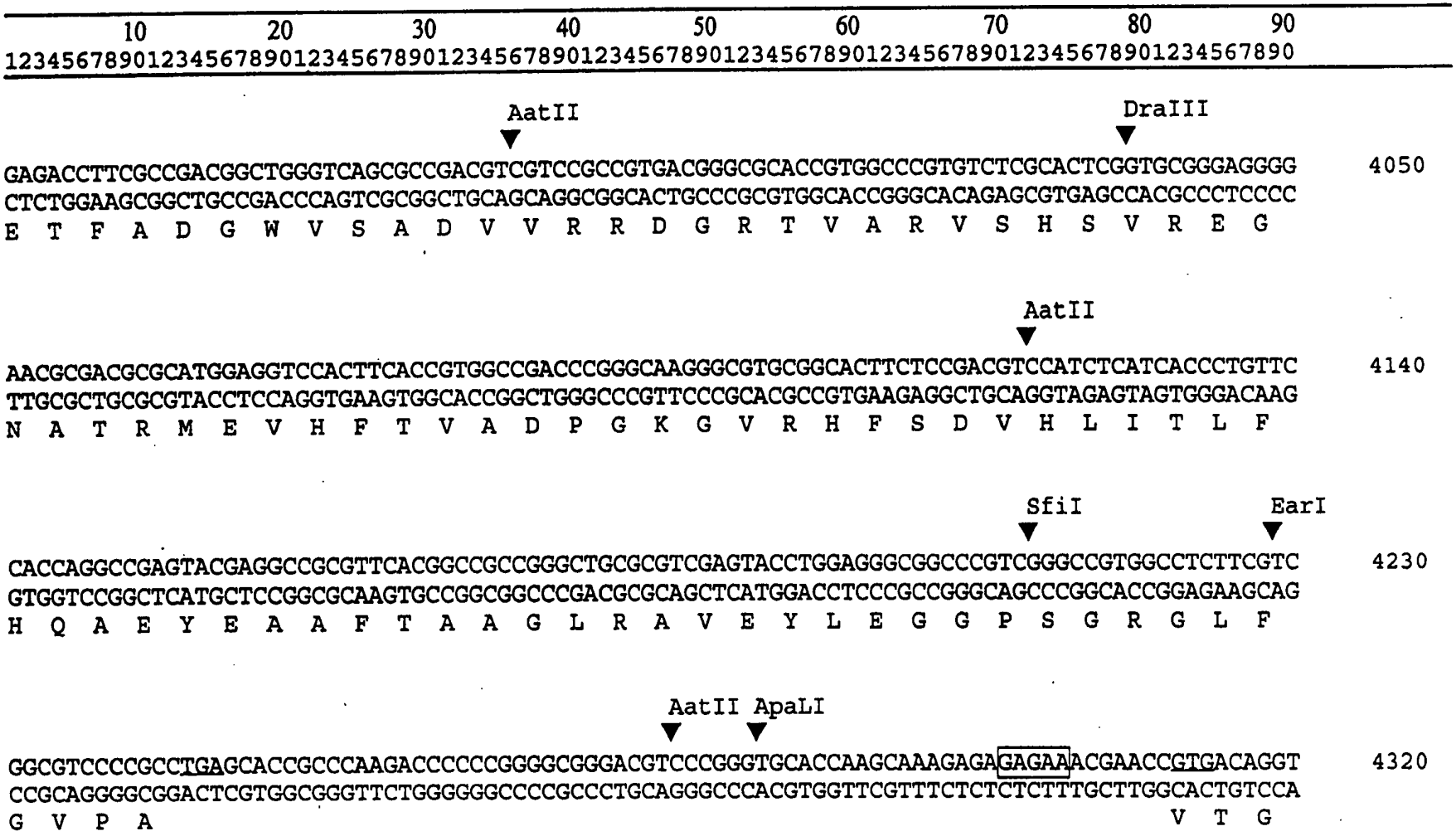

FIG. 32L

sugar.finalgene b-1 Sequence

```
         10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
                                  NotI
                                  ▼
AAGACCCGAATACCGCGTGTCCGCCGCGGCCGCACCACGCCCAGGGCCTTCACCCTGGCCGTCGTCGGCACCCTGCTGGCGGGCACCACC      4410
TTCTGGGCTTATGGCGCACAGGCGGCGCCGGCGTGGTGCGGGTCCCGGAAGTGGGACCGGCAGCAGCCGTGGGACGACCGCCCGTGGTGG
K  T  R  I  P  R  V  R  R  G  R  T  T  P  R  A  F  T  L  A  V  V  G  T  L  L  A  G  T  T

     NotI
     ▼
GTGGCGGCCGCCGCTCCCGGCGCCGCCGACACGGCCAATGTTCAGTACACGAGCCGGGCGGCGGAGCTCGTCGCCCAGATGACGCTCGAC      4500
CACCGCCGGCGGCGAGGGCCGCGGCGGCTGTGCCGGTTACAAGTCATGTGCTCGGCCCGCCGCCTCGAGCAGCGGGTCTACTGCGAGCTG
V  A  A  A  A  P  G  A  A  D  T  A  N  V  Q  Y  T  S  R  A  A  E  L  V  A  Q  M  T  L  D

GAGAAGATCAGCTTCGTCCACTGGGCGCTGGACCCCGACCGGCAGAACGTCGGCTACCTTCCCGGCGTGCCGCGTCTGGGCATCCCGGAG      4590
CTCTTCTAGTCGAAGCAGGTGACCCGCGACCTGGGGCTGGCCGTCTTGCAGCCGATGGAAGGGCCGCACGGCGCAGACCCGTAGGGCCTC
E  K  I  S  F  V  H  W  A  L  D  P  D  R  Q  N  V  G  Y  L  P  G  V  P  R  L  G  I  P  E

                                                                  MscI
                                                                  BalI
                                                                  ▼
CTGCGTGCCGCCGACGGCCCGAACGGCATCCGCCTGGTGGGGCAGACCGCCACCGCGCTGCCCGCGCCGGTCGCCCTGGCCAGCACCTTC      4680
GACGCACGGCGGCTGCCGGGCTTGCCGTAGGCGGACCACCCCGTCTGGCGGTGGCGCGACGGGCGCGGCCAGCGGGACCGGTCGTGGAAG
L  R  A  A  D  G  P  N  G  I  R  L  V  G  Q  T  A  T  A  L  P  A  P  V  A  L  A  S  T  F
```

FIG. 32M

sugar.finalgene b-1 Sequence

```
         10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
```

```
        StyI
        NcoI                                                          PflMI       ApaI
        ▼                                                             ▼           ▼
GACGACACCATGGCCGACAGCTACGGCAAGGTCATGGGCCGCGACGGTCGCGCGCTCAACCAGGACATGGTCCTGGGCCCGATGATGAAC      4770
CTGCTGTGGTACCGGCTGTCGATGCCGTTCCAGTACCCGGCGCTGCCAGCGCGCGAGTTGGTCCTGTACCAGGACCCGGGCTACTACTTG
D  D  T  M  A  D  S  Y  G  K  V  M  G  R  D  G  R  A  L  N  Q  D  M  V  L  G  P  M  M  N

AACATCCGGGTGCCGCACGGCGGCCGGAACTACGAGACCTTCAGCGAGGACCCCCTGGTCTCCTCGCGCACCGCGGTCGCCCAGATCAAG      4860
TTGTAGGCCCACGGCGTGCCGCCGGCCTTGATGCTCTGGAAGTCGCTCCTGGGGGACCAGAGGAGCGCGTGGCGCCAGCGGGTCTAGTTC
N  I  R  V  P  H  G  G  R  N  Y  E  T  F  S  E  D  P  L  V  S  S  R  T  A  V  A  Q  I  K

GGCATCCAGGGTGCGGGTCTGATGACCACGGCCAAGCACTTCGCGGCCAACAACCAGGAGAACAACCGCTTCTCCGTGAACGCCAATGTC      4950
CCGTAGGTCCCACGCCCAGACTACTGGTGCCGGTTCGTGAAGCGCCGGTTGTTGGTCCTCTTGTTGGCGAAGAGGCACTTGCGGTTACAG
G  I  Q  G  A  G  L  M  T  T  A  K  H  F  A  A  N  N  Q  E  N  N  R  F  S  V  N  A  N  V

                                                  StyI        SfiI
                                                  ▼           ▼
GACGAGCAGACGCTCCGCGAGATCGAGTTCCCGGCGTTCGAGGCGTCCTCCAAGGCCGGCGCGGCCTCCTTCATGTGTGCCTACAACGGC      5040
CTGCTCGTCTGCGAGGCGCTCTAGCTCAAGGGCCGCAAGCTCCGCAGGAGGTTCCGGCCGCGCCGGAGGAAGTACACACGGATGTTGCCG
D  E  Q  T  L  R  E  I  E  F  P  A  F  E  A  S  S  K  A  G  A  A  S  F  M  C  A  Y  N  G
```

FIG. 32N

sugar.finalgene b-1 Sequence

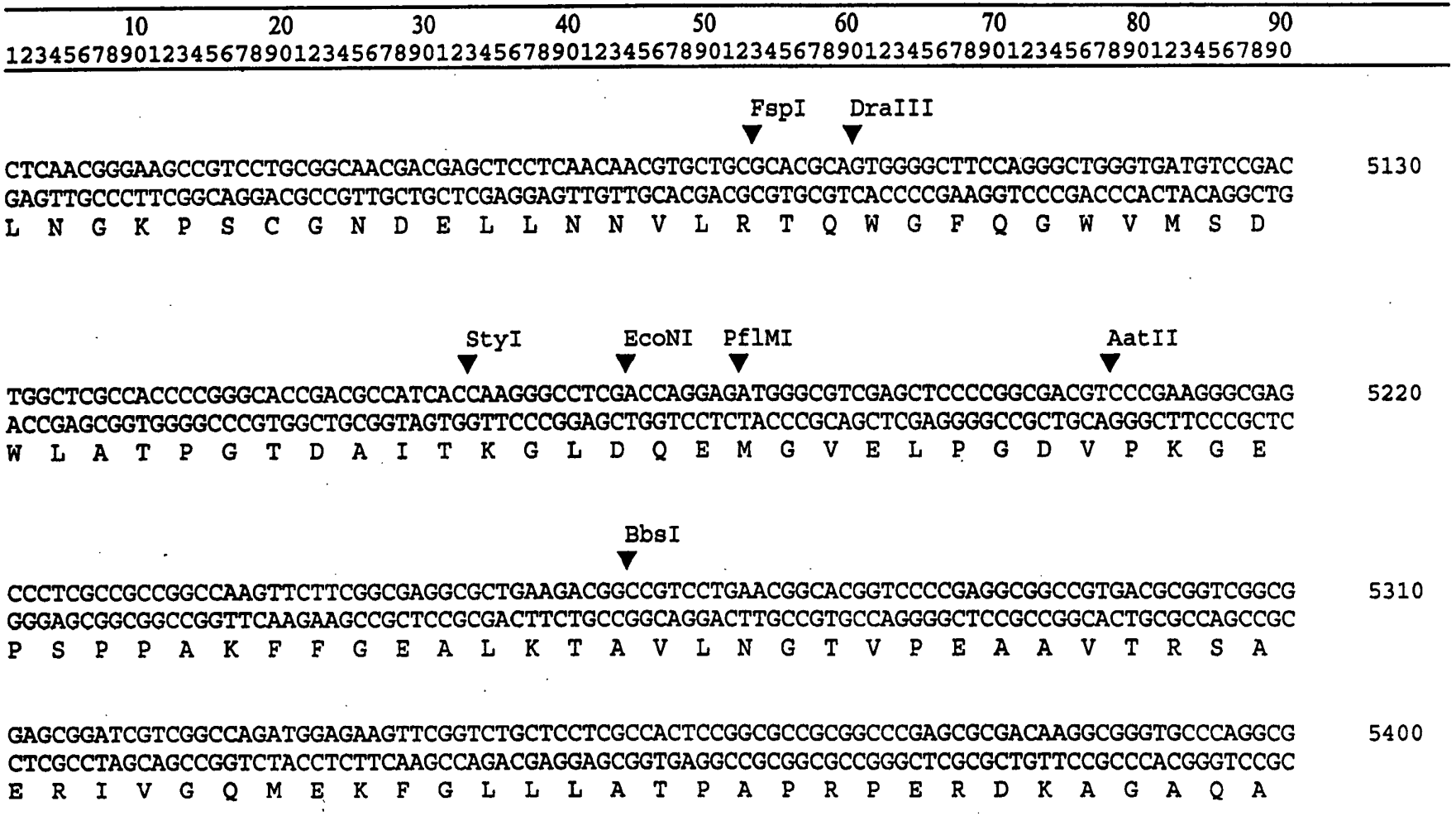

FIG. 320

sugar.finalgene b-1 Sequence

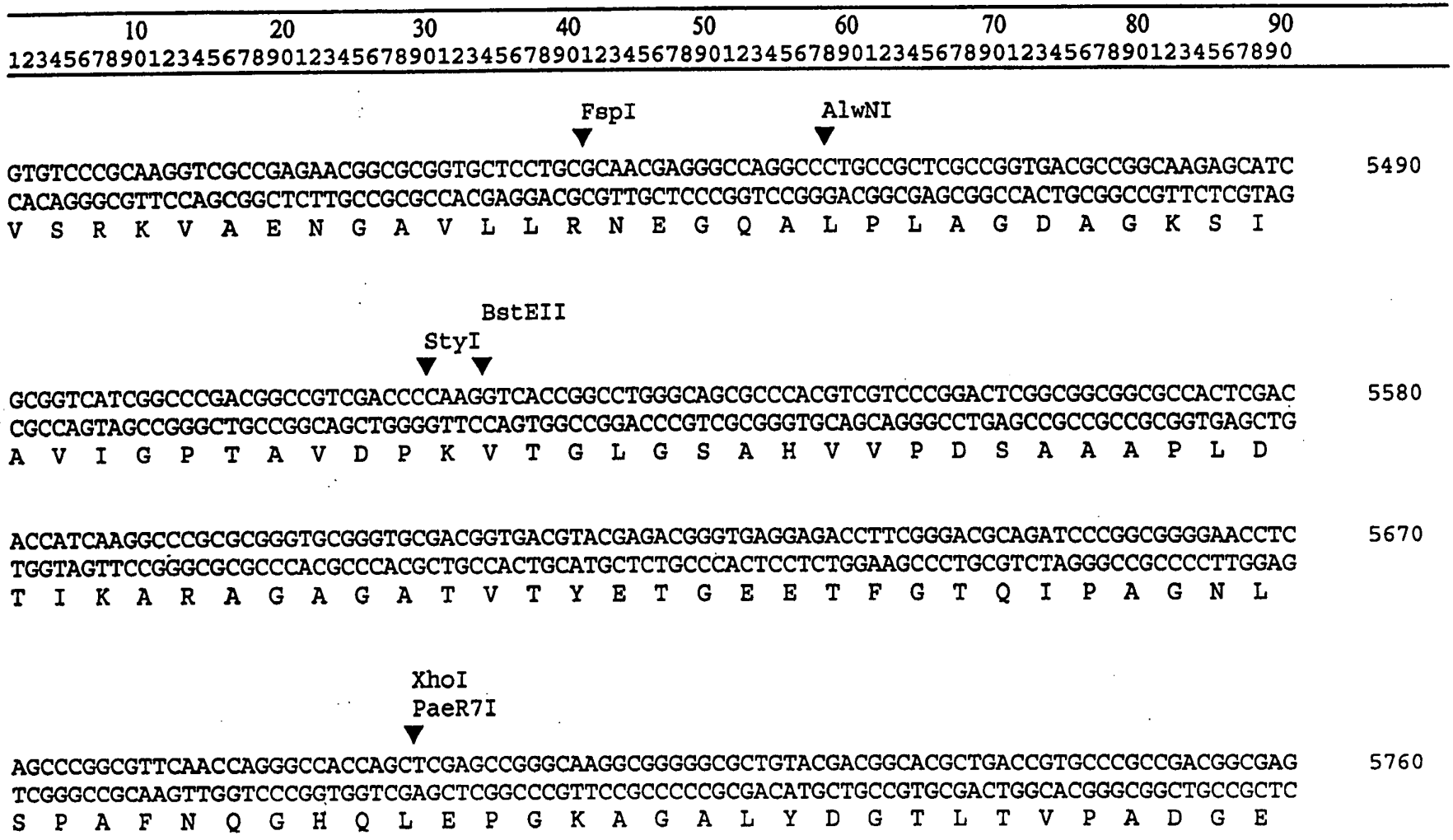

FIG. 32P

sugar.finalgene b-1 Sequence

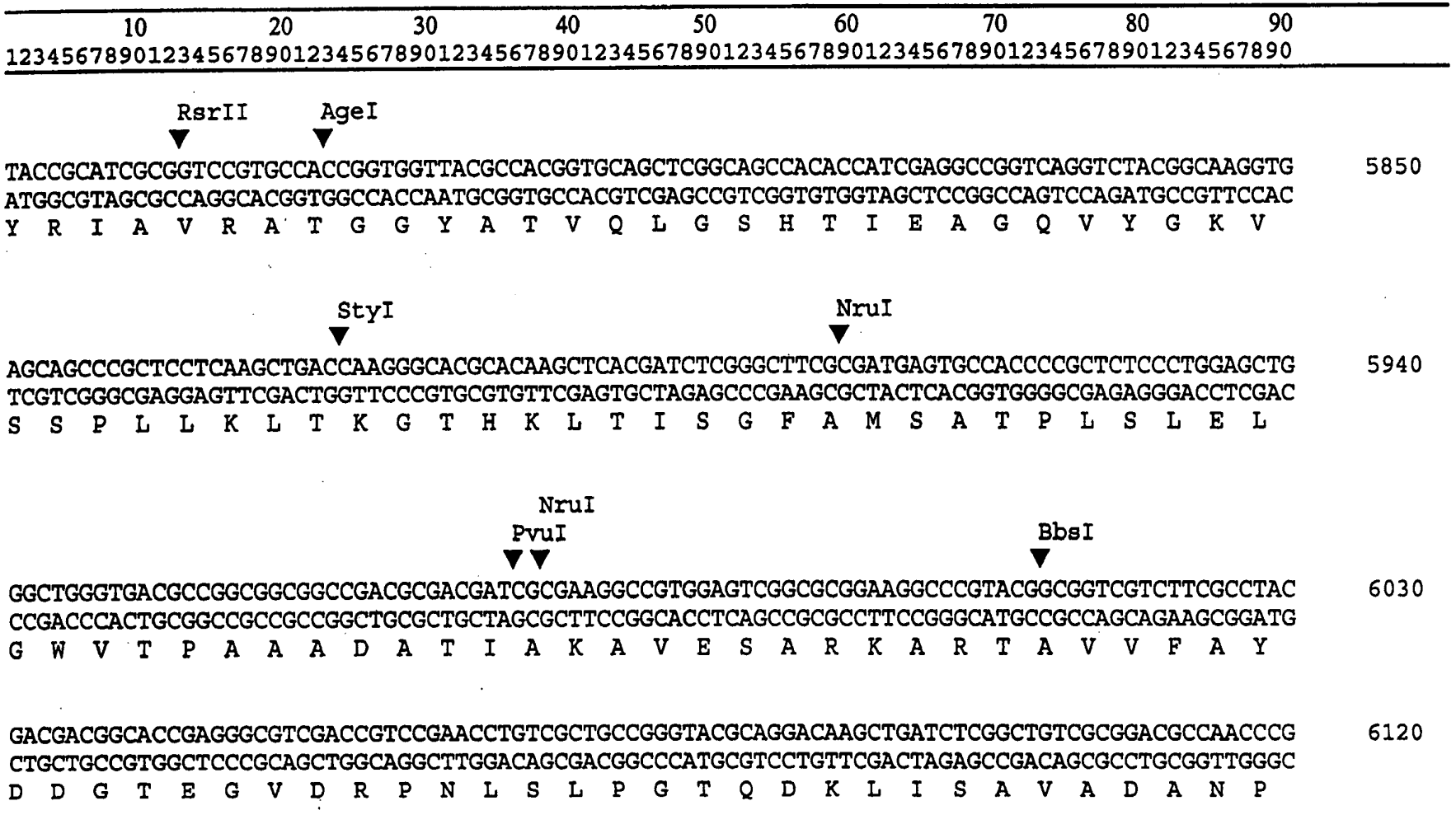

FIG. 32Q

sugar.finalgene b-1 Sequence

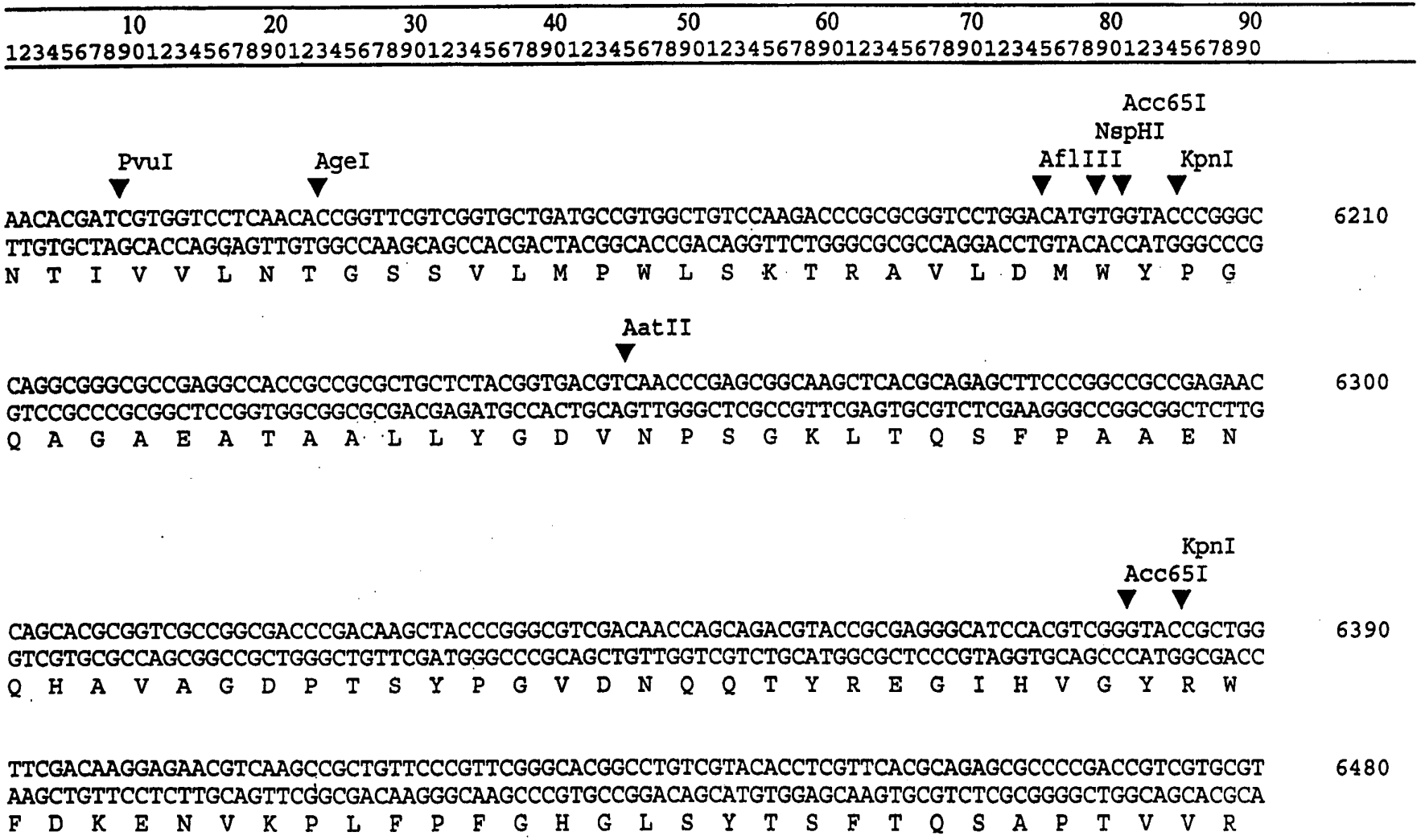

FIG. 32R

sugar.finalgene b-1 Sequence

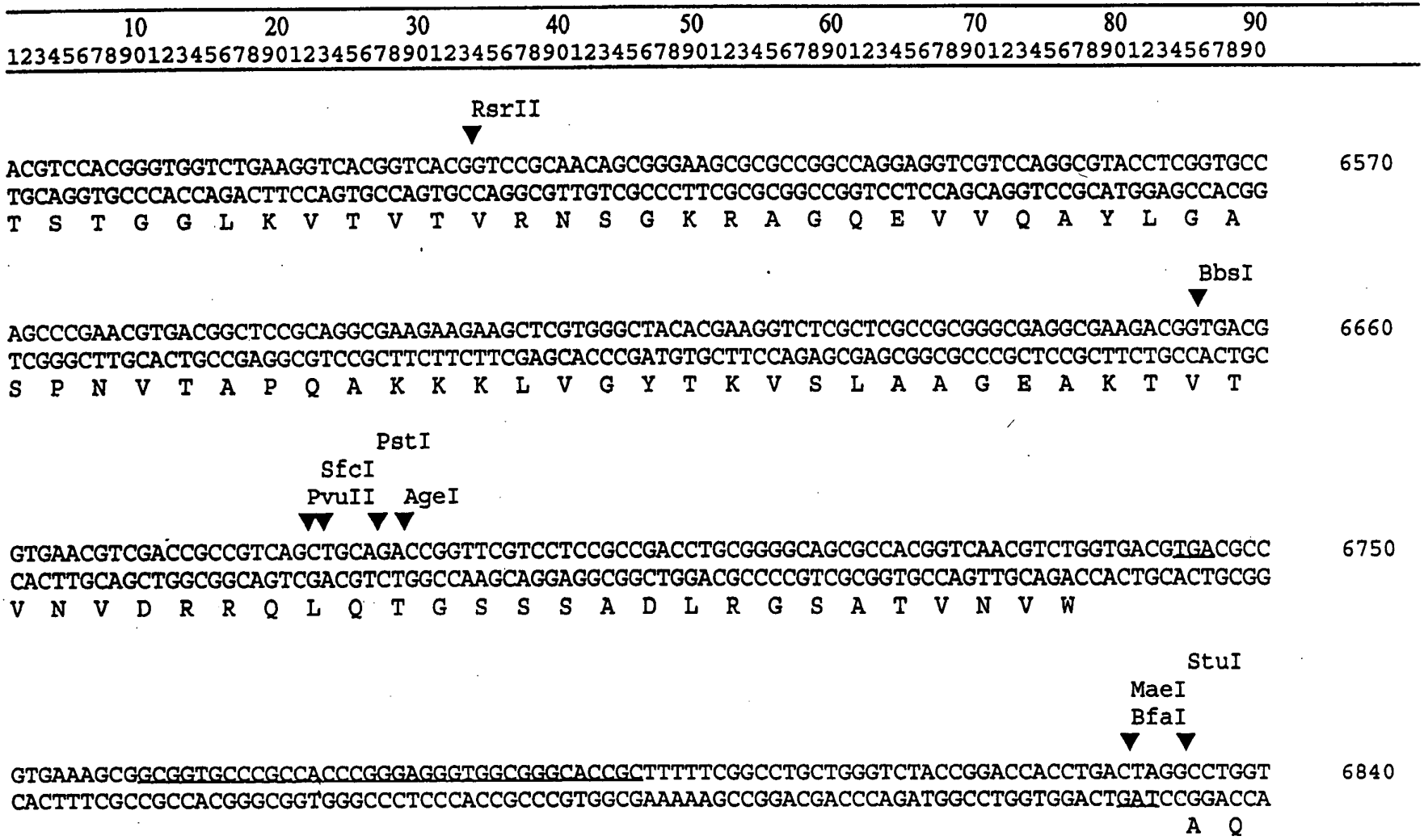

FIG. 32S

sugar.finalgene b-1 Sequence

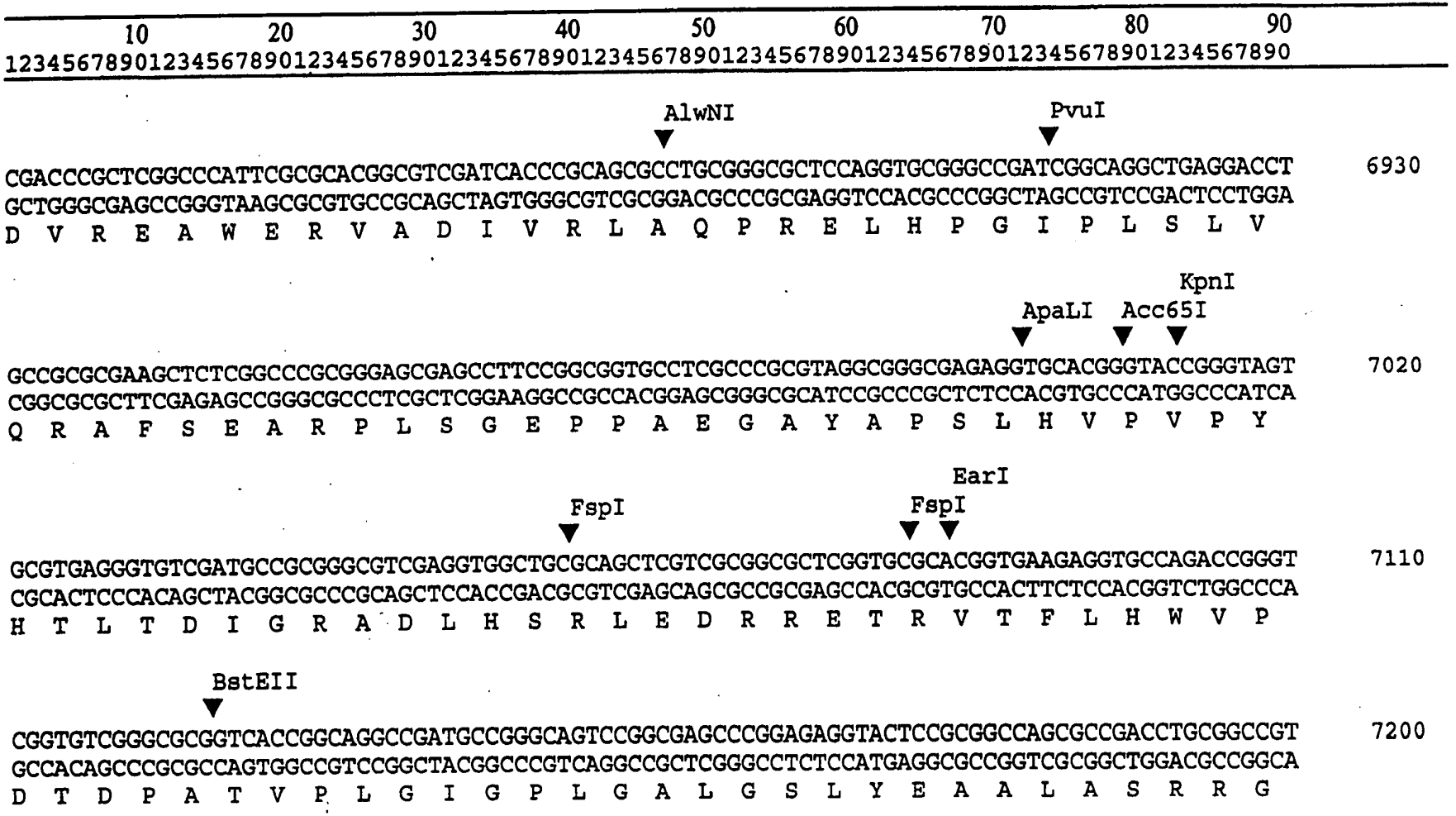

FIG. 32T

sugar.finalgene b-1 Sequence

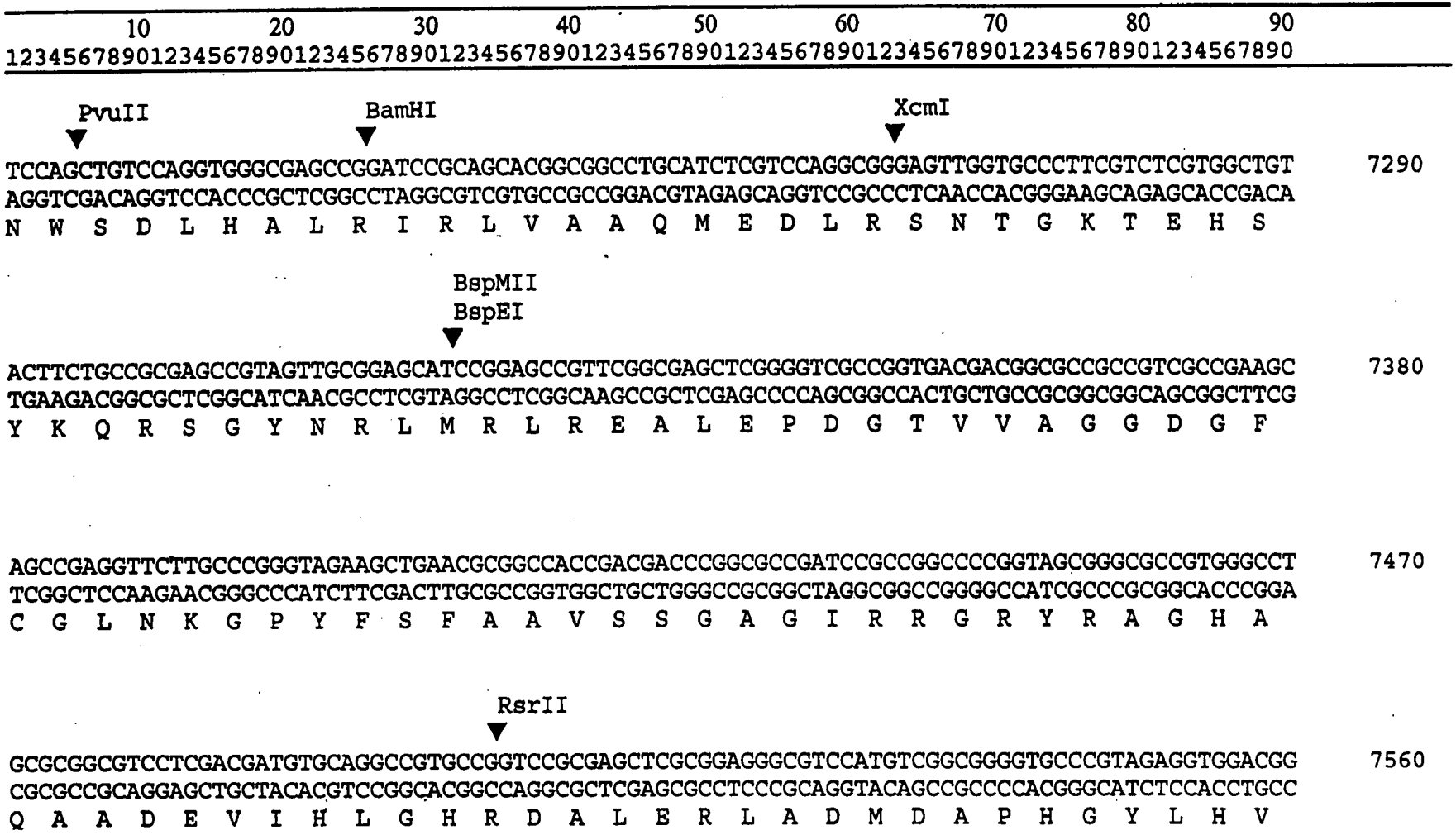

FIG. 32U

sugar.finalgene b-1 Sequence

```
         10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
```

```
GGAGGAGCGCCCGGGTGCGGGGGGTGATCGCCTTCTCGACGAGCAGCGGGTCCAGGGTGGGGTGGTCCTCGTGCGGCTCGACGGGCACGG      7650
CCTCCTCGCGGGCCCACGCCCCCCACTAGCGGAAGAGCTGCTCGTCGCCCAGGTCCCACCCCACCAGGAGCACGCCGAGCTGCCCGTGCC
P  L  L  A  R  T  R  P  T  I  A  K  E  V  L  L  P  D  L  T  P  H  D  E  H  P  E  V  P  V

                                        BsaAI
                          PvuII         AflIII
                            ▼            ▼▼
GGGTCGCGCCGGTGGCGGACACCGCGAGCCAGCTGGCGATGTACGTGTGCGAGGGGACGATCACCTCGTCCCCGGGTCCGATGCCGAGGC      7740
CCCAGCGCGGCCACCGCCTGTGGCGCTCGGTCGACCGCTACATGCACACGCTCCCCTGCTAGTGGAGCAGGGGCCCAGGCTACGGCTCCG
P  T  A  G  T  A  S  V  A  L  W  S  A  I  Y  T  H  S  P  V  I  V  E  D  G  P  G  I  G  L

                                                                                  TfiI
                                                                                   ▼
CGCGGAGGGCGAGCTGGAGGGCGTCCATCCCGCTGTTCACGCCGACGGCGTGGTCCGTCTCGCAGTACGCGGCGAACTCCGCCTCGAATC      7830
GCGCCTCCCGCTCGACCTCCCGCAGGTAGGGCGACAAGTGCGGCTGCCGCACCAGGCAGAGCGTCATGCGCCGCTTGAGGCGGAGCTTAG
G  R  L  A  L  Q  L  A  D  M  G  S  N  V  G  V  A  H  D  T  E  C  Y  A  A  F  E  A  E  F

                                                  PvuI
                                                   ▼
CTTCGAGTTCGGGTCCGAGGAGGTAGCGCCCCGAGTCGAGGACGCGGGCGATCGCGGCGTCGGTCTCCGCGCGGAGCTCCTCGTAGGCGG      7920
GAAGCTCAAGCCCAGGCTCCTCCATCGCGGGGCTCAGCTCCTGCGCCCGCTAGCGCCGCAGCCAGAGGCGCGCCTCGAGGAGCATCCGCC
G  E  L  E  P  G  L  L  Y  R  G  S  D  L  V  R  A  I  A  A  D  T  E  A  R  L  E  E  Y  A
```

FIG. 32V

sugar.finalgene b-1 Sequence

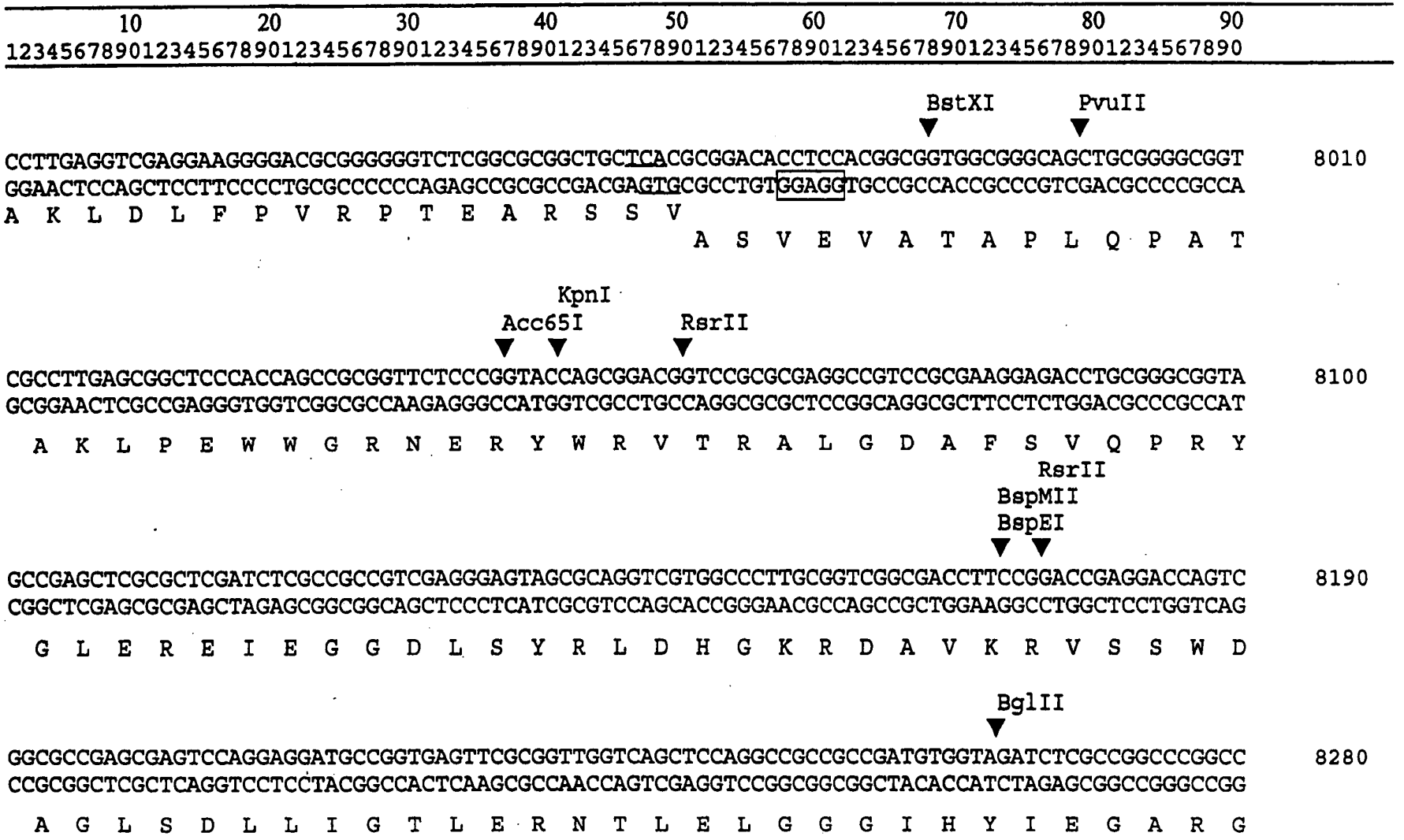

FIG. 32W

sugar.finalgene b-1 Sequence

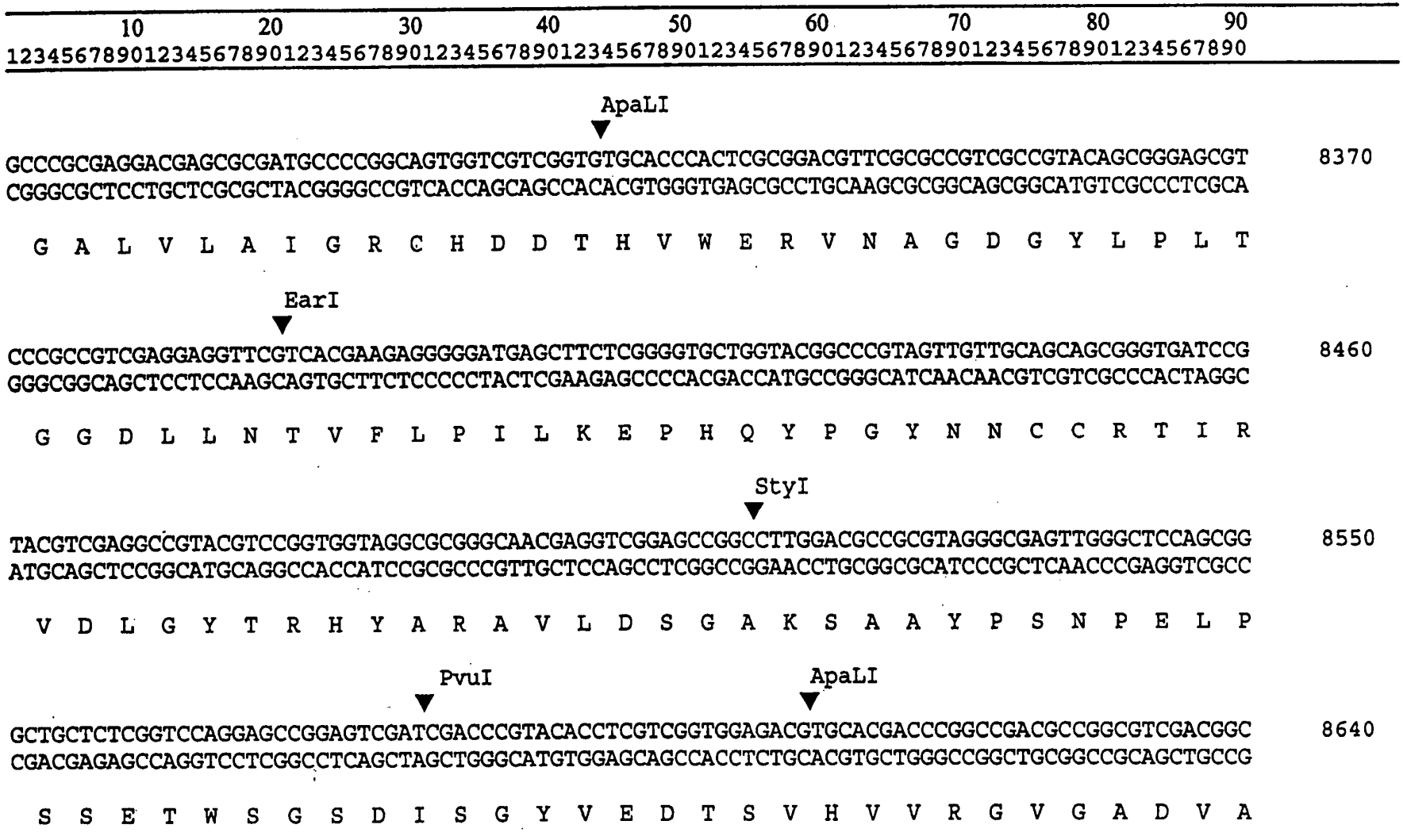

FIG. 32X

sugar.finalgene b-1 Sequence

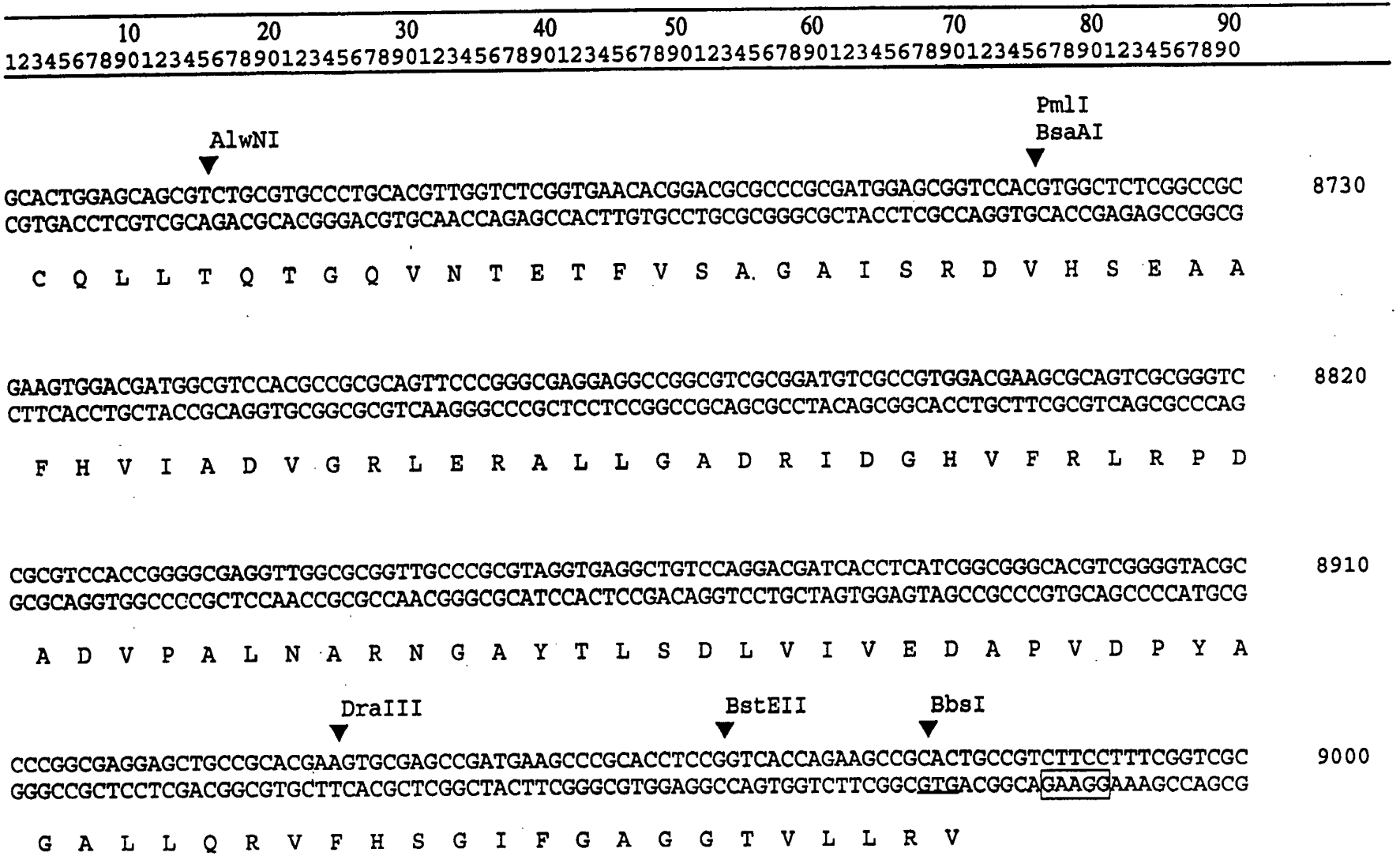

FIG. 32Y

sugar.finalgene b-1 Sequence

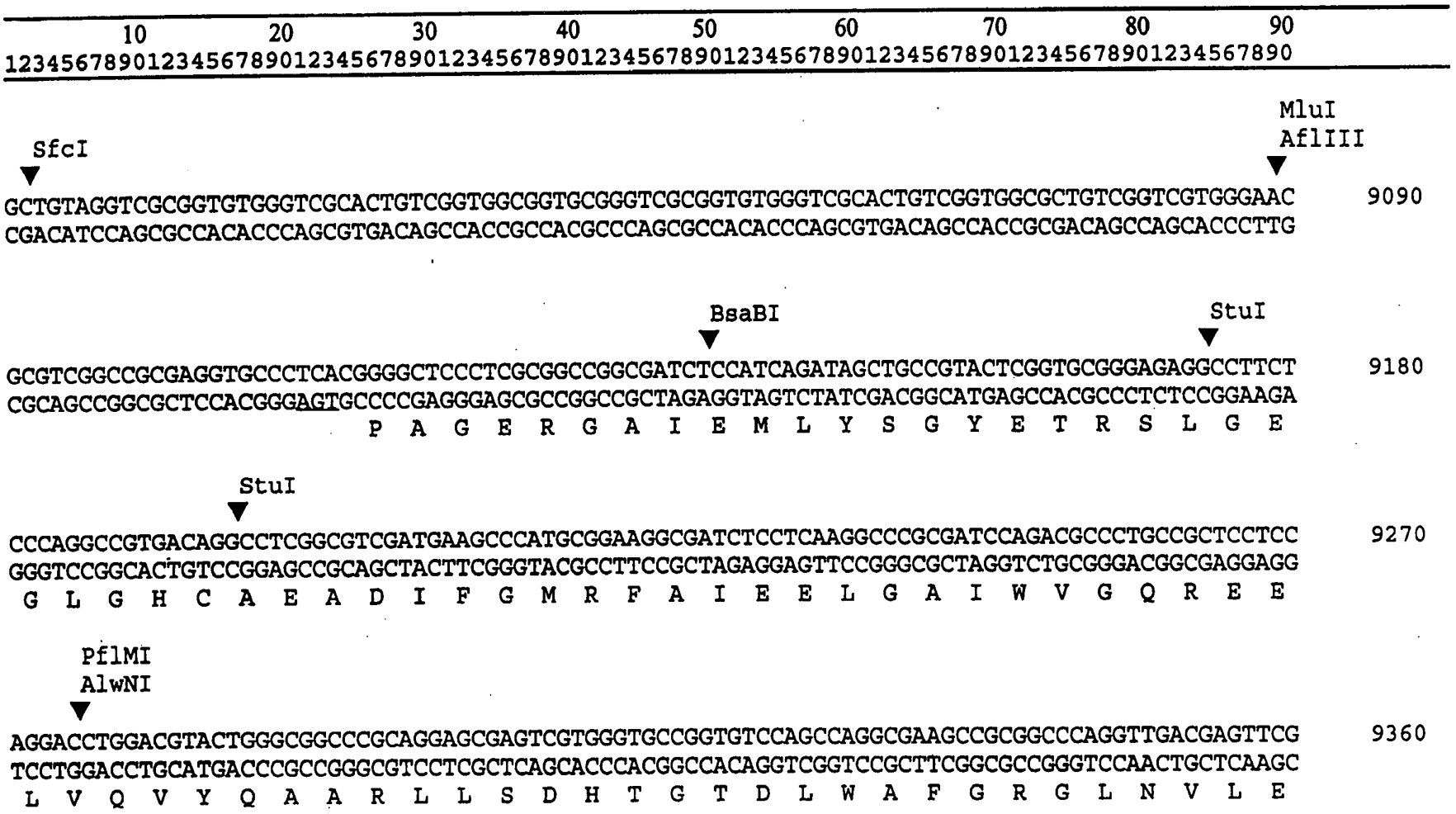

FIG. 32Z

sugar.finalgene b-1 Sequence

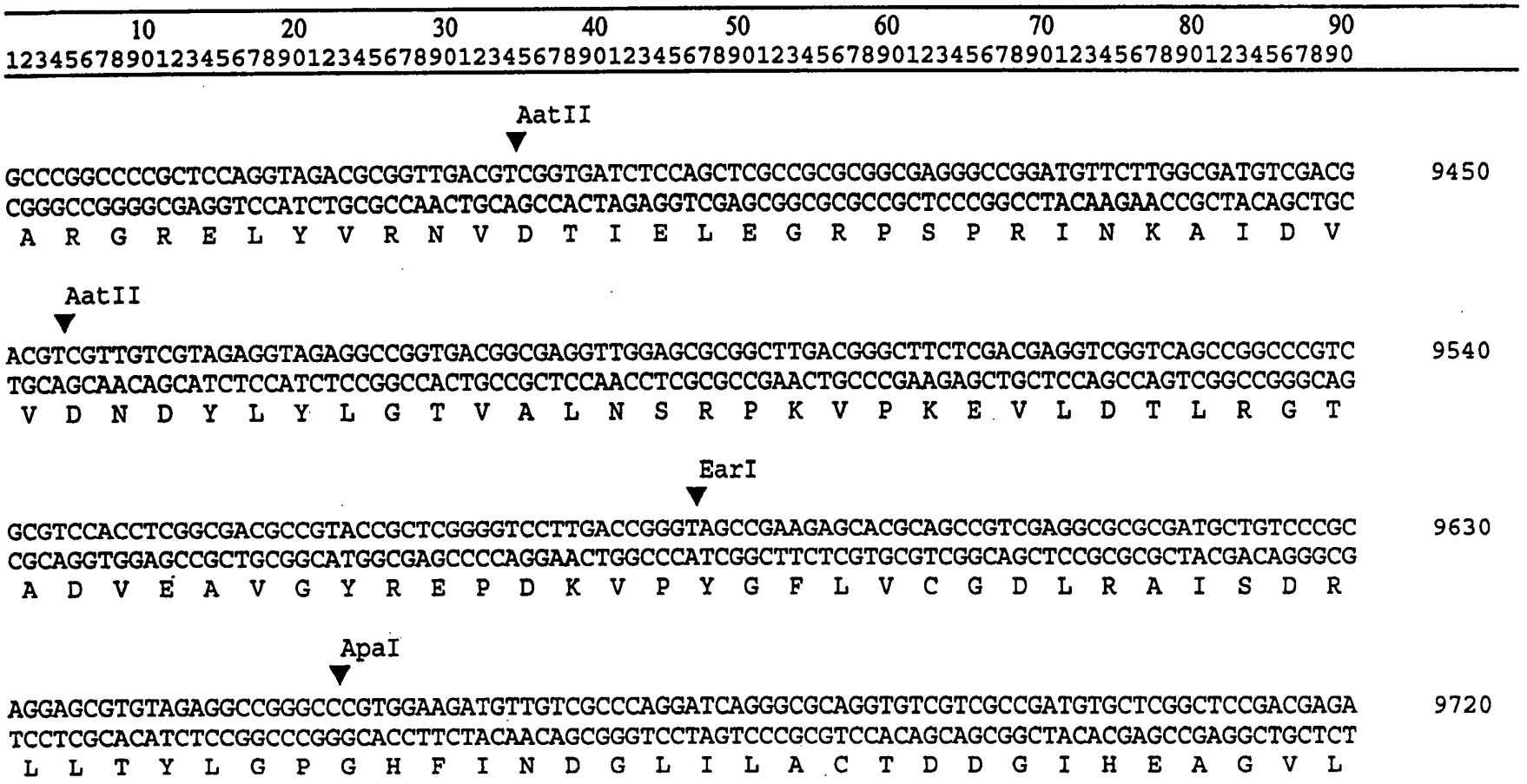

FIG. 32AA

sugar.finalgene b-1 Sequence

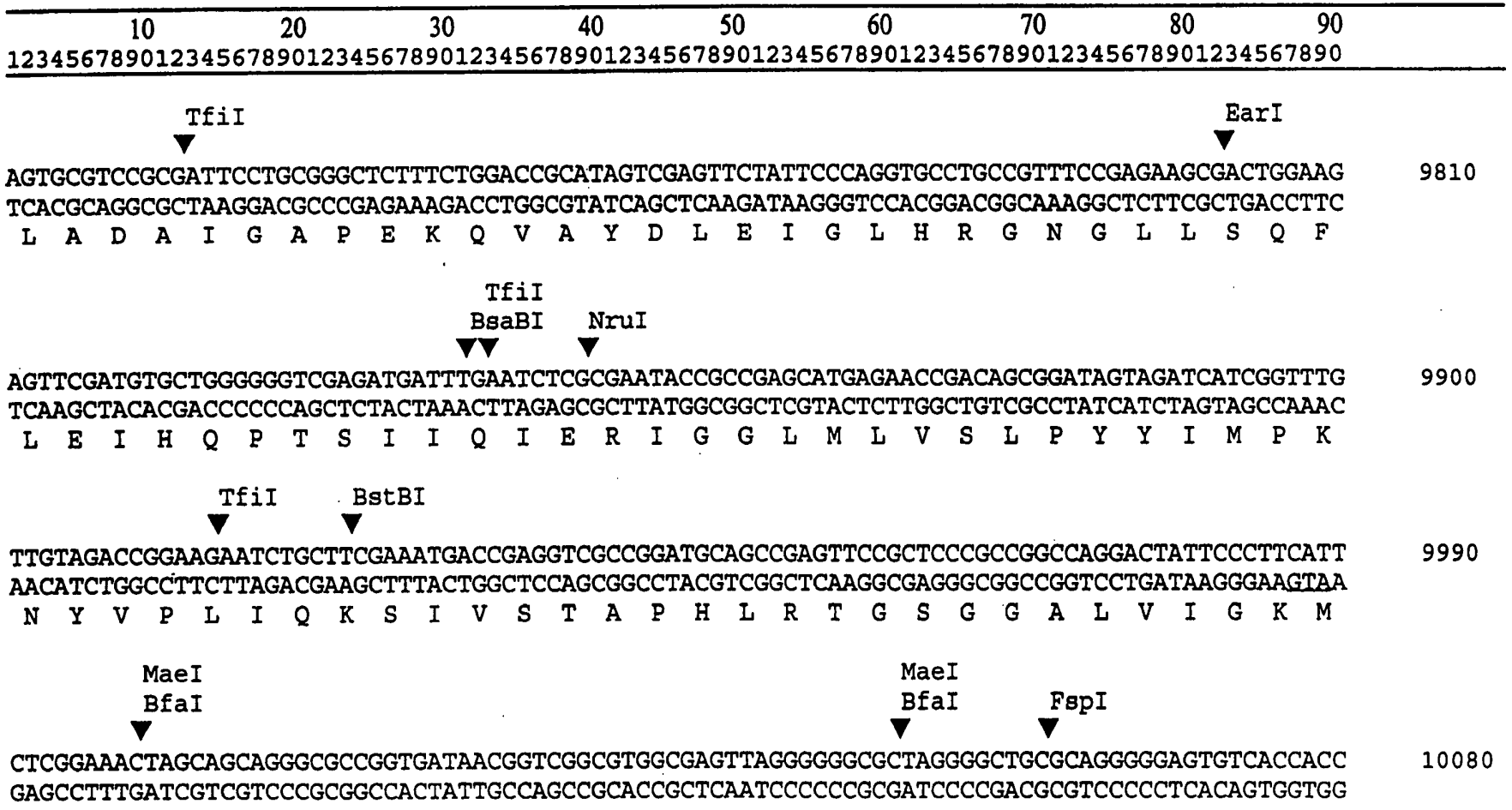

FIG. 32BB

sugar.finalgene b-1 Sequence

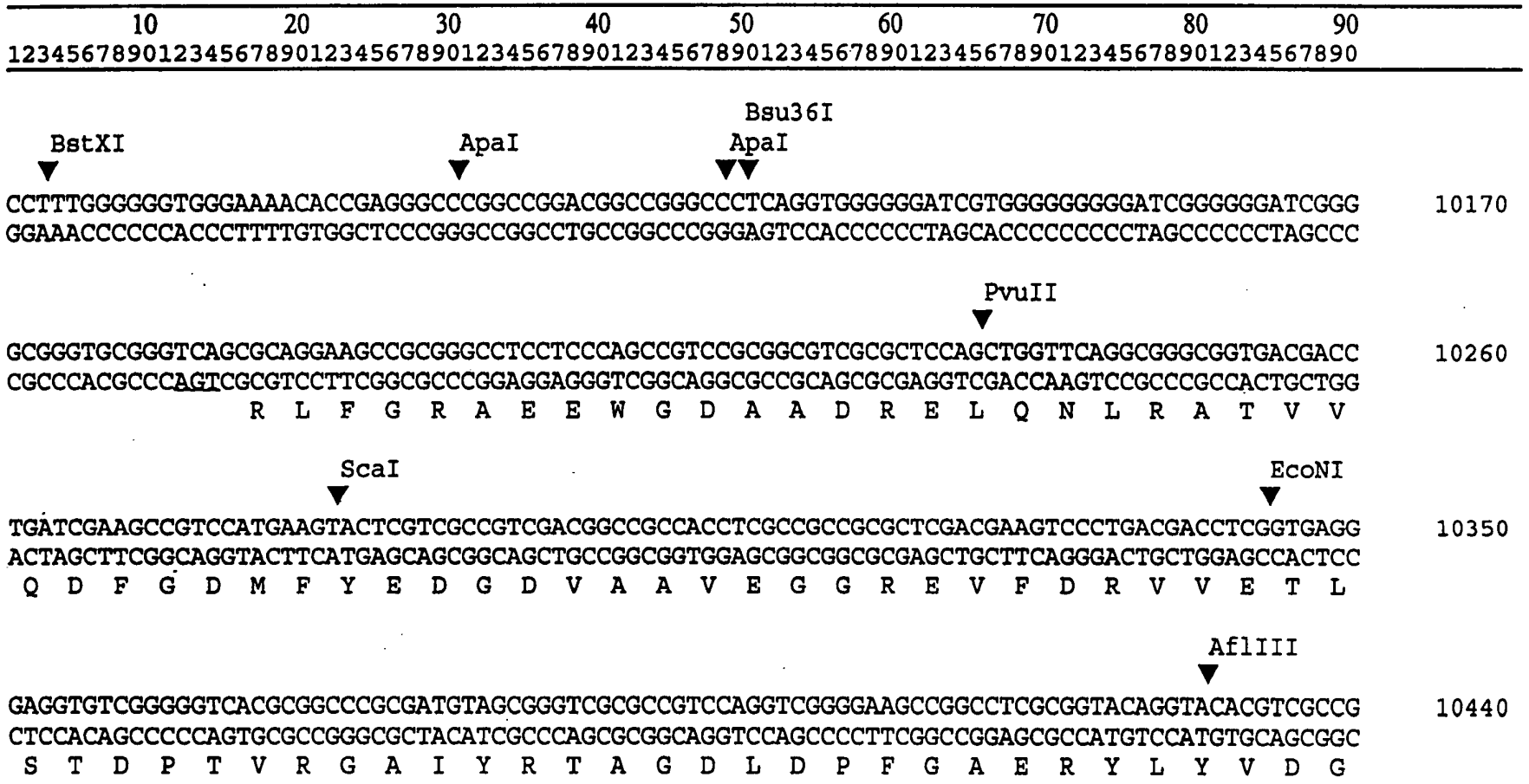

FIG. 32CC

sugar.finalgene b-1 Sequence

```
         10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
```

```
AGGAGATCGACCTGCACCGCGACCTGCGGGTGCGCGGTGGGCCGCATGGTGGCGGGCTTGATCCGCAGCAGTTCGGCGTCGGCCCCGGTG      10530
TCCTCTAGCTGGACGTGGCGCTGGACGCCCACGCGCCACCCGGCGTACCACCGCCCGAACTAGGCGTCGTCAAGCCGCAGCCGGGGCCAC
 L  L  D  V  Q  V  A  V  Q  P  H  A  T  P  R  M  T  A  P  K  I  R  L  L  E  A  D  A  G  T

 FspI
 ▼
CGCAGGCTGTTCAGGGCGTAGCCGTAGTCGATGTGGAGTCCGGGGGTGCGCTCGCGGACCCGCTCCTCGAAGGCGTTGAGGGCCTCCTGG      10620
GCGTCCGACAAGTCCCGCATCGGCATCAGCTACACCTCAGGCCCCCACGCGAGCGCCTGGGCGAGGAGCTTCCGCAACTCCCGGAGGACC
 R  L  S  N  L  A  Y  G  Y  D  I  H  L  G  P  T  R  E  R  V  R  E  E  F  A  N  L  A  E  Q

                                          SfcI        NruI                  PvuI
                                          ▼           ▼                     ▼
AGCTCGGCCCGCTCCTCCTGCGGCAGCTTGCCGTCGTCACGGCCGCTGTAGTCCTCGCGAATGTTGACGAAGTCGATCGTCCTGCCCTGC      10710
TCGAGCCGGGCGAGGAGGACGCCGTCGAACGGCAGCAGTGCCGGCGACATCAGGAGCGCTTACAACTGCTTCAGCTAGCAGGACGGGACG
 L  E  A  R  E  E  Q  P  L  K  G  D  D  R  G  S  Y  D  E  R  I  N  V  F  D  I  T  R  G  Q

CCGGCGTCGTTGAGGTCGGCGATGAAGTCGACCAGGTCGAGCAGGCGGGAGGCACGGCCCGGGAGCACGATGTAGGCGAAGCCGAGGTTG      10800
GGCCGCAGCAACTCCAGCCGCTACTTCAGCTGGTCCAGCTCGTCCGCCCTCCGTGCCGGGCCCTCGTGCTACATCCGCTTCGGCTCCAAC
 G  A  D  N  L  D  A  I  F  D  V  L  D  L  L  R  S  A  R  G  P  L  V  I  Y  A  F  G  L  N
```

FIG. 32DD

sugar.finalgene b-1 Sequence

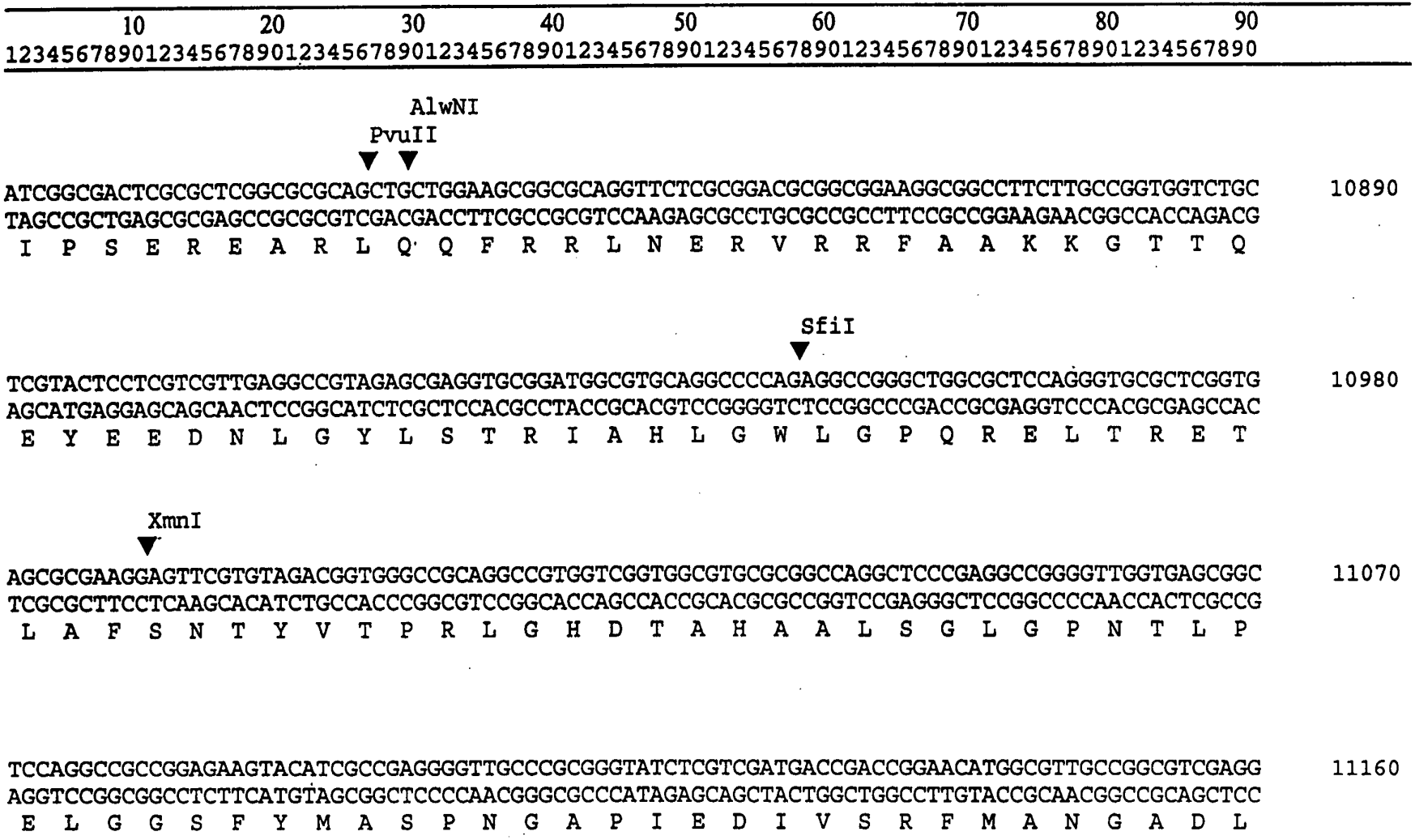

FIG. 32EE

sugar.finalgene b-1 Sequence

```
         10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
                                                    NspHI
                                                    ▼
GCGGACGGGTCGTAGCGGGCGCCGGTCACACGGACGCAGAAGTGGCAGCGGAACATGCAGGTCGGGCCGGGGTAGAGGCCGACGCTGTAC      11250
CGCCTGCCCAGCATCGCCCGCGGCCAGTGTGCCTGCGTCTTCACCGTCGCCTTGTACGTCCAGCCCGGCCCCATCTCCGGCTGCGACATG
 A  S  P  D  Y  R  A  G  T  V  R  V  C  F  H  C  R  F  M  C  T  P  G  P  Y  L  G  V  S  Y

          BbsI                                BbsI
          ▼                                   ▼
GGGAAGACGGGCTTCCTGGCGAGCGCCGCGTCGAAGACGCCGCGCTGTTCGAGCGGGAGCAGGGTGTTCTTCCAGTACGCCCCGGCGGGG      11340
CCCTTCTGCCCGAAGGACCGCTCGCGGCGCAGCTTCTGCGGCGCGACAAGCTCGCCCTCGTCCCACAAGAAGGTCATGCGGGGCCGCCCC
 P  F  V  P  K  R  A  L  A  A  D  F  V  G  R  Q  E  L  P  L  L  T  N  K  W  Y  A  G  A  P

                                                                                PflMI
                                                                                ▼
CCGGTCTCGACCGCGGTGCGGAGCTCCGGGACCTGCCCGAACAGGGCGAGGAGGCGCCGGAAGGCGTCCCGGTCGACGCCCAGGTCGTGG      11430
GGCCAGAGCTGGCGCCACGCCTCGAGGCCCTGGACGGGCTTGTCCCGCTCCTCCGCGGCCTTCCGCAGGGCCAGCTGCGGGTCCAGCACC
 G  T  E  V  A  T  R  L  E  P  V  Q  G  F  L  A  L  L  R  R  F  A  D  R  D  V  G  L  D  H

CGGGCCTCCTCCAGCGGGGTGAAGGGGCTGTTGCCGTAGCGCACGGCGAGCCGGACGAGGTGGCGGGCGGTCGTTCCGGCCTCGTCGGGC      11520
GCCCGGAGGAGGTCGCCCCACTTCCCCGACAACGGCATCGCGTGCCGCTCGGCCTGCTCCACCGCCCGCCAGCAAGGCCGGAGCAGCCCG
 R  A  E  E  L  P  T  F  P  S  N  G  Y  R  V  A  L  R  V  L  H  R  A  T  T  G  A  E  D  P
```

FIG. 32FF

sugar.finalgene b-1 Sequence

```
         10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
```

```
GGCACGAGGCCGCCGGCGGCGAGGGTCTGGCCGACGGCGTGGACCGCCGCCCCCAGATCGGCTCCGGGGTGCGCGCAGCGTTCGGCCGGG      11610
CCGTGCTCCGGCGGCCGCCGCTCCCAGACCGGCTGCCGCACCTGGCGGCGGGGGTCTAGCCGAGGCCCCACGCGCGTCGCAAGCCGGCCC
 P  V  L  G  G  A  A  L  T  Q  G  V  A  H  V  A  A  G  L  D  A  G  P  H  A  C  R  E  A  P

GCGGTGGCGGAAAGGGCGGGGGCGGTCATCGGGAGCGTCCAATCGTGGGCGTGGATGTCTGGGGGGCCGCGAGCGGGGCGGGGGCCGTGT      11700
CGCCACCGCCTTTCCCGCCCCCGCCAGTAGCCCTCGCAGGTTAGCACCCGCACCTACAGACCCCCCGGCGCTCGCCCCGCCCCCGGCACA
 A  T  A  S  L  A  P  A  T  M
                              R  S  R  G  I  T  P  T  S  T  Q  P  A  A  L  P  A  P  A  T

                        NotI
                          ▼
CGCGGTGGCGCGCGGTCAGTTCGCGGCCGCGGGTCGCGCAGAGACGCAGCAGGTCGGCGACCCGGCGGATGTCGTCGTCGCCGATGGCGG      11790
GCGCCACCGCGCGCCAGTCAAGCGCCGGCGCCCAGCGCGTCTCTGCGTCGTCCAGCCGCTGGGCCGCCTACAGCAGCAGCGGCTACCGCC

 D  R  H  R  A  T  L  E  R  G  R  T  A  C  L  R  L  L  D  A  V  R  R  I  D  D  D  G  I  A

                                                                                  PflMI
                                                                                    ▼
TGCCGGTCGGCAGGGACAGCACGCGCGCGGCGAGGCGTTCGGTGTGCGGCAGCGGGGCGTGCGGCTGCCCGCGGTACGGCTCCAGCTCGT      11880
ACGGCCAGCCGTCCCTGTCGTGCGCGCGCCGCTCCGCAAGCCACACGCCGTCGCCCCGCACGCCGACGGGCGCCATGCCGAGGTCGAGCA

 T  G  T  P  L  S  L  V  R  A  A  L  R  E  T  H  P  L  P  A  H  P  Q  G  R  Y  P  E  L  E
```

FIG. 32GG

sugar.finalgene b-1 Sequence

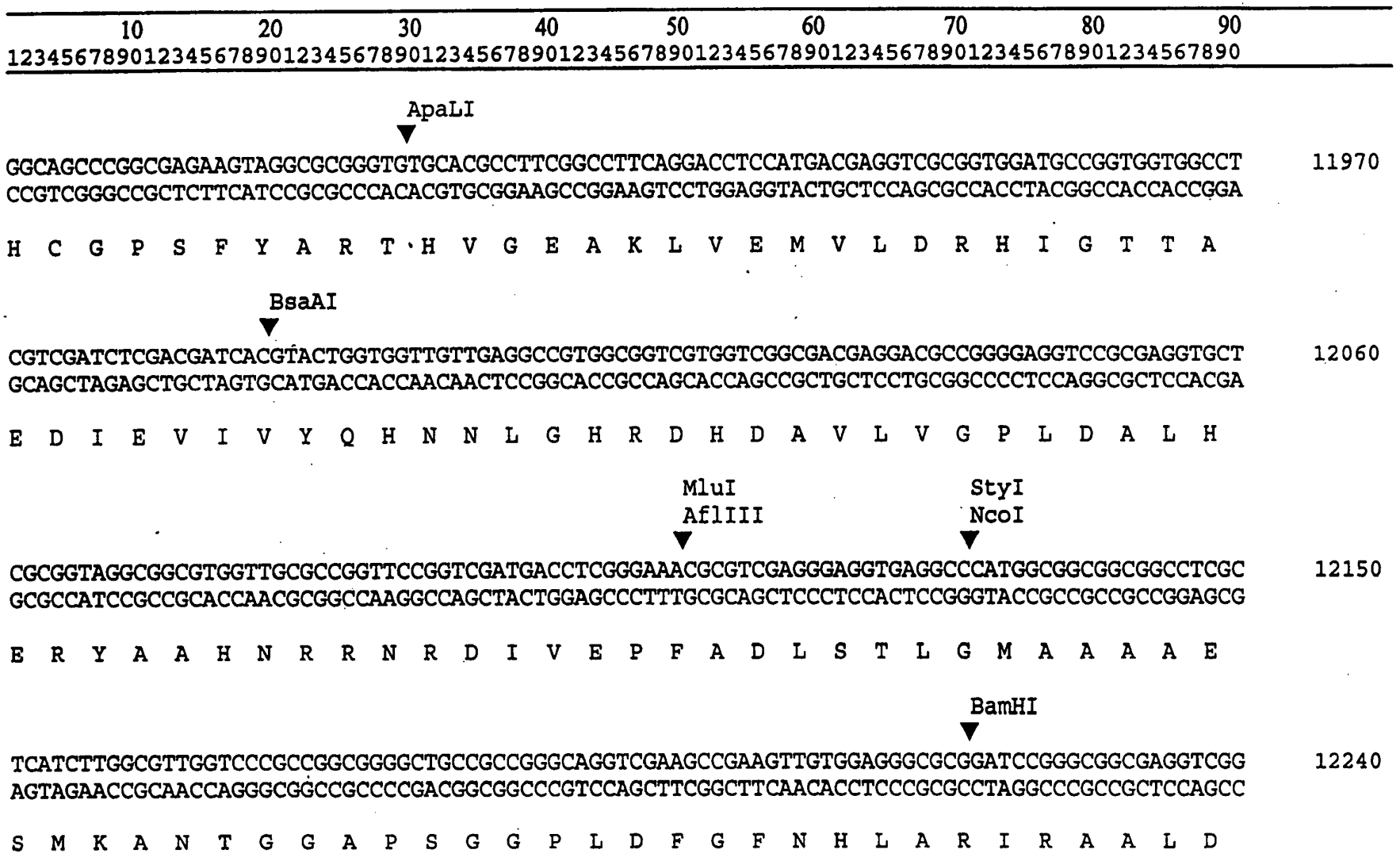

FIG. 32HH

sugar.finalgene b-1 Sequence

```
         10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
                                          StyI                      BbsI
                                           ▼                         ▼
CGTCGTCGGTGACGACGGCGCCGCCCTCGAAGGCGTTGACGGCCTTGGTGGCGTGGAAGCTGAAGACCTCGGCGTCGCCGAGGCTGCCGG      12330
GCAGCAGCCACTGCTGCCGCGGCGGGAGCTTCCGCAACTGCCGGAACCACCGCACCTTCGACTTCTGGAGCCGCAGCGGCTCCGACGGCC

A  D  D  T  V  V  A  G  G  E  F  A  N  V  A  K  T  A  H  F  S  F  V  E  A  D  G  L  S  G

                                                                                PvuII
                                                                                  ▼
CGGGCCGGCCGTCGACCGCGCAGCCGAGGGCGTGCGCGGCGTCGAAGTACAGCCGCAGGCCGTGCTCGTCGGCGACCTTCCGCAGCTGGT      12420
GCCCGGCCGGCAGCTGGCGCGTCGGCTCCCGCACGCGCCGCAGCTTCATGTCGGCGTCCGGCACGAGCAGCCGCTGGAAGGCGTCGACCA

A  P  R  G  D  V  A  C  G  L  A  H  A  A  D  F  Y  L  R  L  G  H  E  D  A  V  K  R  L  Q


CGGCGGCGCAGGGGCGGCCCCAGAGGTGGACGCCGACGACGGCCGAGGTGCGGGGTGTGACCGCGGCGGCCACCTGGTCCGGGTCGAGGT      12510
GCCGCCGCGTCCCCGCCGGGGTCTCCACCTGCGGCTGCTGCCGGCTCCACGCCCCACACTGGCGCCGCCGGTGGACCAGGCCCAGCTCCA

D  A  A  C  P  R  G  W  L  H  V  G  V  V  A  S  T  R  P  T  V  A  A  A  V  Q  D  P  D  L

                                  BbsI
                                   ▼
TGCCGGTGTCCGGGTCGATGTCGGCGAAGACCGGGGTGAGGCCGATCCAGCGCAGTGCGTGCGGGGTGGCGGCGAACGTCATCGACGGCA      12600
ACGGCCACAGGCCCAGCTACAGCCGCTTCTGGCCCCACTCCGGCTAGGTCGCGTCACGCACGCCCCACCGCCGCTTGCAGTAGCTGCCGT

N  G  T  D  P  D  I  D  A  F  V  P  T  L  G  I  W  R  L  A  H  P  T  A  A  F  T  M  S  P
```

FIG. 32II

sugar.finalgene b-1 Sequence

```
         10        20        30        40        50        60        70        80        90
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789
```

```
                                                                    MscI      SphI
 BclI                                                               BalI      NspHI
 ▼                                                                  ▼         ▼
TGATCACTTCGCCGGTGAGGCCGGCGGCGTGCGCGAGGAGCTGGAGCCCGGCCGTGGCGTTGCAGGTGGCCACGGCATGCCGGACCCCGG      12690
ACTAGTGAAGCGGCCACTCCGGCCGCCGCACGCGCTCCTCGACCTCGGGCCGGCACCGCAACGTCCACCGGTGCCGTACGGCCTGGGGCC

M  I  V  E  G  T  L  G  A  A  H  A  L  L  Q  L  G  A  T  A  N  C  T  A  V  A  H  R  V  G

                                                          AlwNI          ApaI
                                                          ▼              ▼
CGAGCCCGGCGACGCGCTCCTCGAACTCGCGGACGAGCGGGCCGCCGTTGGACAGCCACTGGCTGTCGAGGGCCCGGTCGAGCCGCTCGT      12780
GCTCGGGCCGCTGCGCGAGGAGCTTGAGCGCCTGCTCGCCCGGCGGCAACCTGTCGGTGACCGACAGCTCCCGGGCCAGCTCGGCGAGCA

A  L  G  A  V  R  E  E  F  E  R  V  L  P  G  G  N  S  L  W  Q  S  D  L  A  R  D  L  R  E

                                                                    BsmI
                                                                    ▼
ACAGCCTGGCGCGGTCGATGCGGTTGGGCCGCCCCACGAGGAGCGGCTGGTCGAAAGCGGCGGGGCCGCCGAAGAATGCGAGGTCGGATA      12870
TGTCGGACCGCGCCAGCTACGCCAACCCGGCGGGGTGCTCCTCGCCGACCAGCTTTCGCCGCCCCGGCGGCTTCTTACGCTCCAGCCTAT

Y  L  R  A  R  D  I  R  N  P  R  G  V  L  L  P  Q  D  F  A  A  P  G  G  F  F  A  L  D  S

                                                TfiI
                                                XmnI                TfiI
                                                ▼▼                  ▼
AGGCGCTTTTCACGGATGTTCCCTCCGGGCCACCGTCACGAAATGATTCGCCGATCCGGGAATCCCGAACGAGGTCGCCGCGCTCCACCG      12960
TCCGCGAAAAGTGCCTACAAGGGAGGCCCGGTGGCAGTGCTTTACTAAGCGGCTAGGCCCTTAGGGCTTGCTCCAGCGGCGCGAGGTGGC

L  A  S  K  V
```

FIG. 32JJ

sugar.finalgene b-1 Sequence

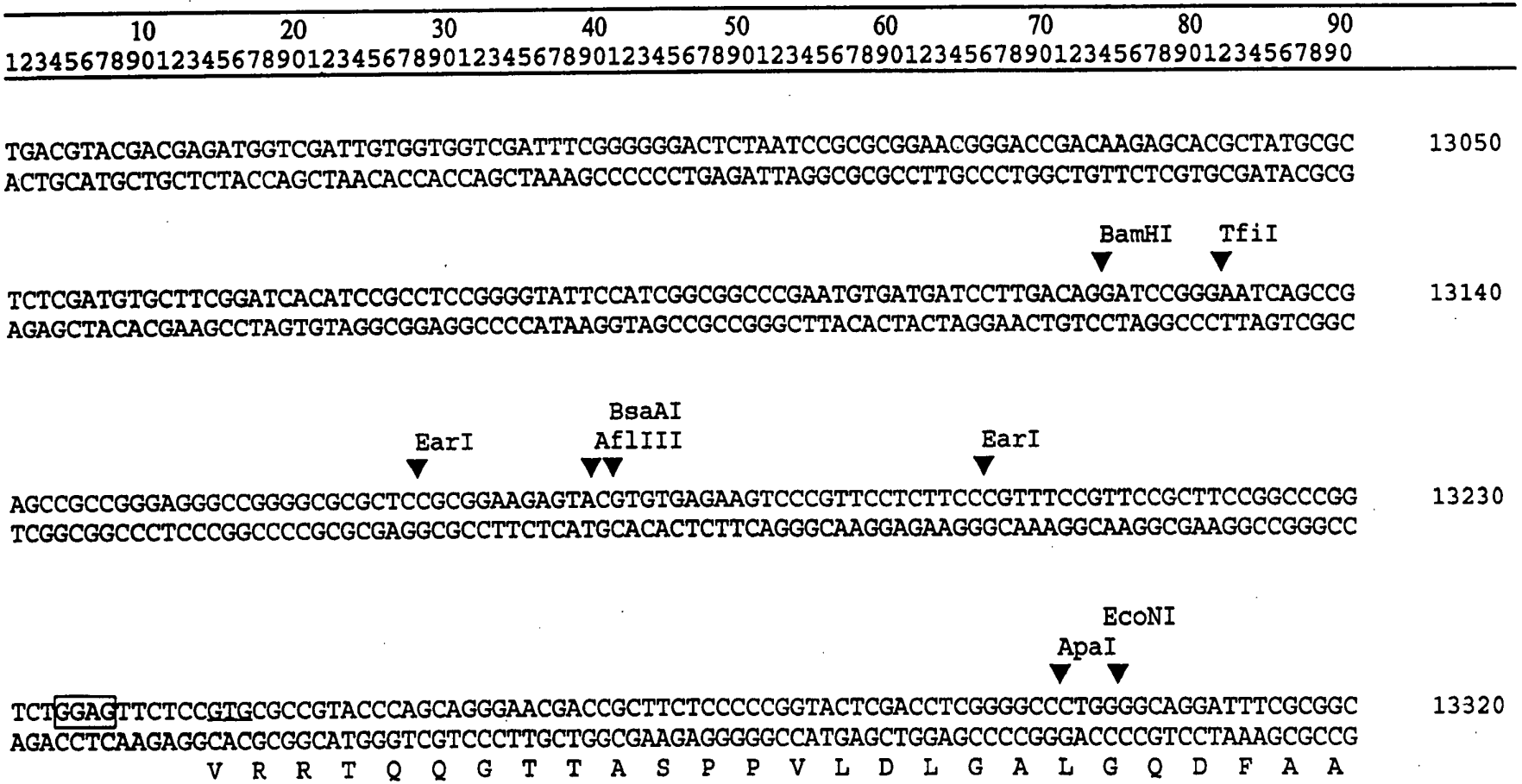

FIG. 32KK

sugar.finalgene b-1 Sequence

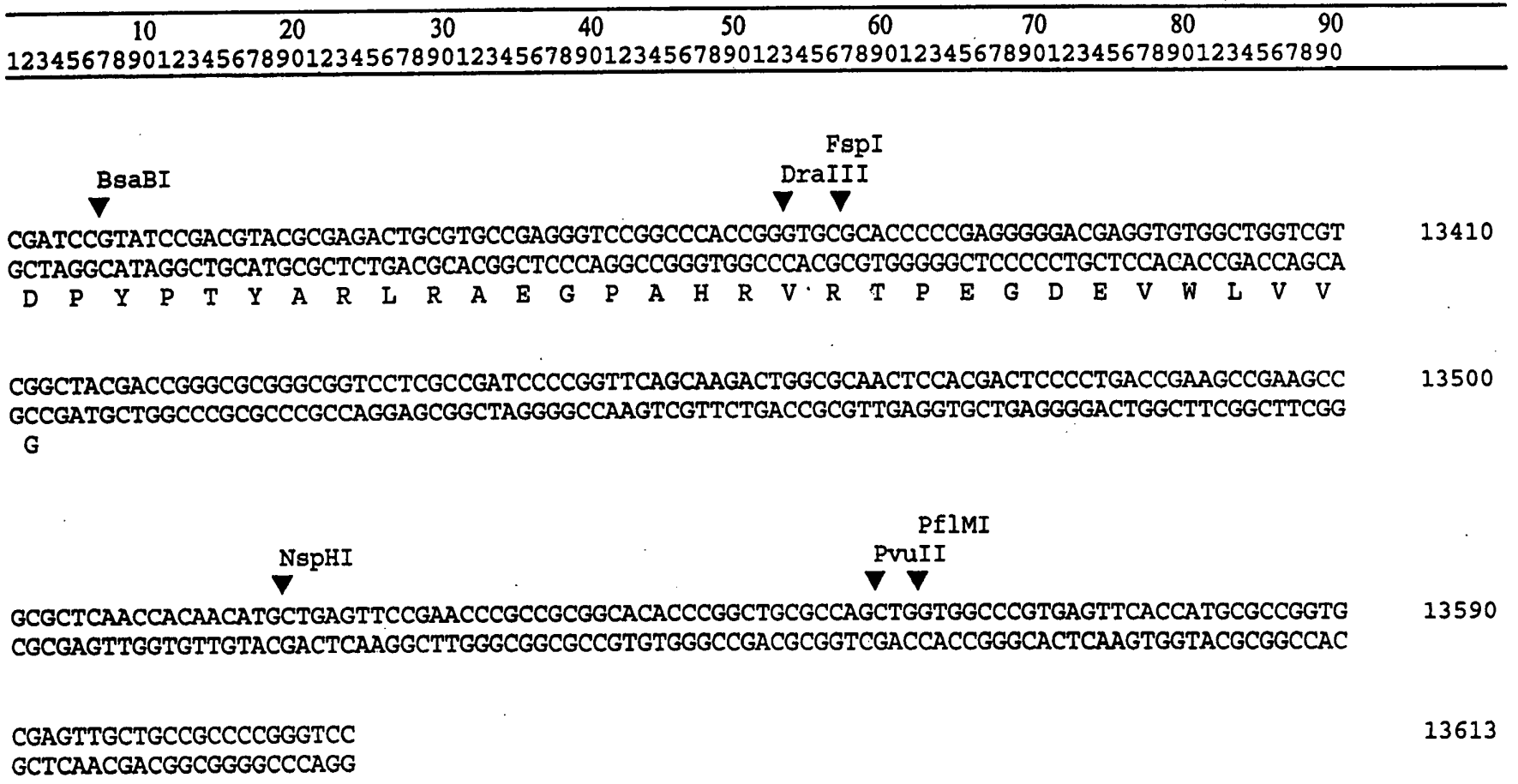

FIG. 32LL